Optimal Signaling and Labelling For Constellation-Constrained Communication Systems

by

Hojat Abdolanezhad

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Most communication systems use a finite signal set as their alphabet set to form a codebook to transmit data over a communication channel in a reliable fashion. The problem with Conventional methods which implement Coded Modulation (CM) schemes such as Trellis Coded Modulation or Multi-level Coding Multi-Stage Decoding is their complexity of dealing with codes with different rates in each level which makes the design and implementation a difficult task.

One simple way to implement Coded Modulation is Bit-Interleaved Coded Modulation (BICM) which uses only a single binary encoder to transmit data. Although BICM is a suboptimal scheme compared to CM, its simplicity, from a practical point of view, is a great motivation to design BICM scheme achieving rates close to those obtained by CM. Lots of efforts have been taken place in the past twenty years to design optimal constellation for different snr regimes in CM under various constraints. Some of them are revisited in this study. A novel approach, called Adjustable Weights Model (AWM), will be presented to design constellations which work very well in both CM and BICM schemes. The model also induces a particular labelling on the constellation.

In this work, some properties of AWM are studied. AWM is used to facilitate design of near optimal signalling for CM and BICM schemes. An optimization problem is formed to find the optimal parameters of the proposed model. Global optimization methods are used to solve the optimization problems. It is shown that the optimal points are always on the boundary of the domain by using data processing inequality . Some suboptimal solutions are provided by moment and cumulants matching techniques. The model has the ability to produce different constellations by adjusting its weights. It is well established that the optimal constellation for high snr region is equillattice. This model also converges to an equillattice constellation in high snr region. Number of nonzero weight parameters in the model can vary according to snr, help us to circumvent the saturation problem with conventional CM scheme.

BICM capacity is presented and its relation with CM capacity is discussed. BICM capacity, as a function of snr, is expanded around zero snr. Different constellations and their labeling can be characterized based on the coefficients in the Taylor expansion. The most important difference between CM and BICM is the effect of labelling in the former scheme. Labelling is irrelevant in CM, but greatly influences the system performance in BICM. Effect of labelling and how to search for optimal labeling is part of this study. It is shown that AWM is optimal at medium and low snr regimes. The model coupled with its underlying labelling is first order optimal. Although Gray labelling is optimal at high snr, it is not optimal in the low snr regime. Higher order optimal constellations are defined to be the constellations that have more than one coefficient in their Taylor expansion matched with CM capacity coefficients. It gives us a powerful tool to study constellation in medium snr regime which has not been already discovered. In addition, optimality criterion is provided.

Acknowledgements

Foremost, I would like to express my sincere gratitude to my supervisor Professor Amir Keyvan Khandani for the continuous support of my study and research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

My sincere thanks also goes to Dr. Kamyar Moshksar for his deep knowledge and insightful helps during my research.

I would like to thank my friends Ehsan Seifi, Babak Mamandipour and Ershad Banijamali for the happy time we spent together in the last two years, the time with laughter and mutual encouragement.

I also like to thank my fellows in CST group, Pooya Mahboubi and Akbar Ghasemi for the stimulating discussions and useful advices through my work.

Last but not the least, I would like to take this opportunity to express the profound gratitude from my deep heart to my parents and my siblings for their love and continuous support throughout my life.

To my beloved parents

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Previous Works

An extensive amount of work has been done so far to determine the capacity maximizing modulation for various number of signal points and different dimensionality. Since the Gaussian distribution achieves the capacity of AWGN channel, it was commonly accepted that for approaching channel capacity by a discrete input distribution, the constellation points should be chosen with a Gaussian-like probability distribution. However, [1] showed that it is possible to approach channel capacity with an equi-probable constellation if we choose the location of the points to represent a geometrically Gaussian-like distribution.

Although, from practical implementation perspective and complexity issues, one usually considers the uniform distribution, $P_X(x) = \frac{1}{|\mathcal{X}|}$. This was one of the main motivation of us working with the model introduced and developed through this study. Different researchers were focusing to design constellation to optimize different measure like minimum distance of constellation [2] [3] [4] [5] or BER for uncoded system [6]. [7] studies capacity miximizing constellation for both CM and BICM schemes. It shows that any probabilistic shaping gain can also be obtained by geometric shaping gain.

One somehow surprising observation made by Ungerboeck in his well-known work [8] was that to obtain rates close to capacity of a Gaussian channel with capacity C, it is enough to use a constellation with only 2^{C+1} points. This observation has been improved in [9] to show that 2^{C-1} input levels suffices to achieve rate of C-1 bits and with 2^{C+1} levels we can achieve a rate of C - 0.4.

It was first mentioned in [3] that channel capacity can not be achieved if we use equi-probable equally spaced constellation. [3] discusses the gap between channel capacity and the rate achieved by equi-probable equally spaced one-dimensional constellations. This gap is equal to 1.53 dB for large signal-to-noise ratio [3] [4]. In [9], using bounding techniques, it is shown that in the equi-probable equally spaced case with N points, the capacity approaches $C - \frac{1}{2}\log\frac{\pi e}{6}$ as $N \to \infty$. The gap is smaller for lower snr values [3].

The fundamental idea of coded modulation is to maximize the overall performance of the system by jointly designing of encoder and modulator together. The first breakthrough in implementing CM scheme was introducing TCM (trellis-codeded modulation) by Ungerboeck [8] and MLC (Multilevel coding) by Imai and Hirakawa [10]. Mapping by set partitioning is the main approach of Ungerboeck to Coded Modulation. Set partitioning is based on maximizing the minimum intra-subset Euclidean distance. Imai's approach to Coded Modulation is to transmit each bit position of binary representation of constellation points (labelling) using a binary code C_i at level i. At receiver, each bit level is decode separately, considering it is provided with the already decoded bits in the previous stages. The advantage of MLC is that any code can be used as component code to encode individual levels. High performance codes such as LDPC or Turbo codes have great attraction to be used in this method.

One interesting fact about MLC is its optimality under usage of binary codes. Huber ert al. [11] [12] [13] proved independent of each other that the Coded Modulation capacity can be achieved by multilevel codes in conjunction with Multi-Stage decoding if and only if the rates of the codes in each level is adjusted properly. [10] is a very good reference on theoretical background and also practical rules for designing and implementing coded modulation schemes. It shows how to achieve power and bandwidth-efficient MLC schemes close to theoretical limits.

In BICM scheme, introducing a bit interleaver between the binary encoder and the modulator gives us the benefits that the encoder and modulator become independent from each other. In this case, both of them can be designed separately to optimize the overall performance of the system. For this model, Caire et al. in [14] defined a BICM capacity given by:

$$C_{\mathcal{X}}^{BICM} = \sum_{i=1}^{m} I(B_i; Y) \tag{1.1}$$

and in [15] using typical sequences, the authors show that it is indeed achievable even when the interleaver is not ideal (using a finite-length interleaver). In this case, the choice of labelling highly affects the performance of the system and the resulting rates. For a given constellation, one tries to find the optimum binary labelling from a BICM capacity maximization point of view. [14] conjectured that Gray lebeling maximizes BICM capacity but we show that it is not true for all range of snr. [8] compared Gray labeling with Set Partitioning labeling which is used in TCM (Trellis Coded Modulation) scheme for various PSK and QAM constellations.

In [16] optimal binary labeling, input distributions and input alphabets for BICM capacity are discussed at low snr regime. By some examples it shows that for different snr regions the BICM capacity is maximized with different labelings.

1.2 Contributions

Main contribution of this study is to present a new method for generating capacity approaching constellations which also induces a specific binary labelling on the constellation points. We discuss the advantages of using this model in CM context and also we show that the produced constellation in addition with its specific labelling achieves higher BICM capacity in low and medium snr regions compared to Gray labelling.

The constellations are generated based on the model: $X = \sum_{i=1}^{m} w_i B_i$ which is a weighted sum of m binary random variables taking values from $\{-1, 1\}$. We call our model AWM (Adjustable Weights Model). If $B_i's$ are iid with $P_{B_i}(0) = P_{B_i}(1) = \frac{1}{2}$, then the resulting constellation has 2^m points with probability of choosing each point from the signal set $\mathcal{X}$ the same and equal to $P_X(x_i) = \frac{1}{2^m} \quad i = 1, ..., 2^m$.

We will see that this model characterizes large size constellations with smaller number of parameters, hence, facilitate the optimization of the constellation to make the gap to coded modulation capacity smaller. We will provide several methods to find optimal parameters of the model in an efficient and computationally less expensive ways compared to other constellation optimization algorithms.

AWM is very powerful in the sense that it can adjust its weights to utilize geometric shaping gain in any signal-to-noise ratio. A fixed-size constellation can not achieve rates higher than $\log_2 M$ even for very large snr. The model has the flexibility to increase the constellation size to achieve higher rates with introducing new non-zero weight parameters.

It also shows promising performance to close the gap between BICM capacity and corresponding Coded Modulation capacity at low snr regime. We show the constellation produced by this model is actually First Order Optimal (FOO). As we will see, this model outperforms Gray labelling in some snr region and achieves Gaussian capacity at low snr. Higher order optimality is also studied based on the coefficients of Taylor expansion for BICM capacity around zero snr. An optimality criterion is given to help us characterize constellation and compare performance of different constellation and their associate labelling in low and medium snr.

1.3 Organization

We introduce our proposed model in the second chapter. Some interesting properties of the model is discussed there.

In the third chapter, Coded Modulation scheme is revisited. Different problems in optimal design in Coded Modulation is discussed and an Optimization problem is formulated to find the optimal parameters for CM when generating the input constellation based on our model. Some methods for solving the optimization problem are presented. Among which are genetic algorithm and moment matching.

In chapter four, a more practical approach to coded modulation is discussed; Bit-Interleaved Coded Modulation. In BICM, labelling comes to play. Performance of different signalling and their labelling are compared and optimal signalling are characterized based on first and second coefficients in the Taylor expansion of mutual information.

Chapter 5 summarizes the thesis and also suggests some related problem suitable for future research studies.

Chapter 2

Adjustable Weights Model

2.1 Background

Channel capacity, a fundamental concept in information theory, first was introduced by Shannon [17] as the maximum rate which can be transmitted reliably over a communication channel. Following his work, lots of effort have been taken place to obtain expressions for channel capacity of different communication systems models. Mutual information between two random variables X and Y is defined [18]

$$I(X;Y) = E_{X,Y}[\log \frac{P_{Y|X}(Y|X)}{P_Y(Y)}] \tag{2.1}$$

in which expectation is taken with respect to $P_{X,Y}(X,Y)$.

If Y is related to X through a channel model $P_{Y|X}(Y|X)$, the channel capacity is defined [17] [18] as

$$C = \sup_{P_X \in \mathcal{D}} I(X;Y) \tag{2.2}$$

Where the maximization is performed over all the input distributions satisfying the underlying constraints of the problem.

Although the optimal input distribution may be continues in some channel models, [19] shows that it is discrete for most of the known channel models or we can achieve rates very close to channel capacity by a discrete distribution.

Coded Modulation capacity is defined as the maximum mutual information achieved when a constellation is used as the codebook alphabet i.e. elements of the codewords are chosen form a set with finite cardinality $|\mathcal{X}| < \infty$. Suppose X is a random variable with probability P_X over $\mathcal{X}$. The optimization is over the set of alphabet set $\mathcal{X}$ and the probability distributions P_X on $\mathcal{X}$ which satisfy the assumed constraints (average power or peak-power constraints) called $\mathcal{D}$. Hence for a given m, the problem is to find the location and probability of the points for an m-point constellation that gives the maximum rate. Coded Modulation capacity is defined as:

$$C^{CM} = \sup_{\substack{|\mathcal{X}| \leq m \\ P_X \in \mathcal{D}}} I(X;Y) \tag{2.3}$$

Through out this study we assume an Additive White Gaussian Noise (AWGN) channel model and also average power is the only constraint on the system. Hence, (2.3) can be written as

$$C^{CM} = \sup_{\substack{|\mathcal{X}| \leq m \\ E[X^2] \leq P}} I(X;Y) \tag{2.4}$$

2.2 The Proposed Model

In this section, we introduce our proposed model. In this model the input random variable has a special structure:

$$X = \sum_{i=1}^{m} w_i B_i \qquad w_i \in R \quad i = 1, ..., m \tag{2.5}$$

where B_i's are iid Bernouli random variables on $\{-1, 1\}$ with $P(-1) = P(1) = \frac{1}{2}$. With this definition, X is discrete random variable defined over a constellation $\mathcal{X}$. The number of points in the signal set $\mathcal{X}$ is a power of two, 2^m, which is characterized only by m parameters: $w_1, ..., w_m$. The random variable X can be seen as the weighted sum of m Bernouli random variables. We call $\mathbf{w} = (w_1, ..., w_m)$ weight vector or parameter vector interchangeably through this study. Each constellation point has probability equal to $\frac{1}{2^m}$ unless some of them happens to occur at the same point of real line. In this case, the probabilities of points add up and the resulting constellation is no longer equi-probable. Another fact about our model is that, the resulting constellation is always symmetric, regardless of whether they overlap or not.

We can generalize this model to produce high-dimensional constellations. Suppose C is a $M \times m$ matrix with $M = 2^m$ and entries come from $\{-1, 1\}$. Each row of C is a vertex of a 2^m-dimesnional zero-mean hyper-cube. now

$$X = C\ W \tag{2.6}$$

represents a 2^m-point constellation in N-dimensional space for W a $m \times N$ matrix of real numbers. X is an $M \times N$ matrix in which each row represents a constellation point. This model can represents any constellation if one can solve the linear system in (2.6) for W.

Suppose we want to see if this model can produce an equi-probable equi-distannt 16-QAM constellation. In this case M=16 and $m = \log_2 M = 4$ and

$$C = \begin{bmatrix} -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{2.7}$$

We assume constellation points to form a $M \times N$ matrix, each point corresponds to a row of X.

$$X = \begin{bmatrix} -3 & 3 \\ -1 & 3 \\ \vdots & \vdots \\ 3 & -3 \end{bmatrix} \tag{2.8}$$

solving (2.6) for W gives

$$W = \begin{bmatrix} -0.0000 & 2.0000 \\ -0.0000 & 1.0000 \\ 2.0000 & 0.0000 \\ 1.0000 & 0.0000 \end{bmatrix} \tag{2.9}$$

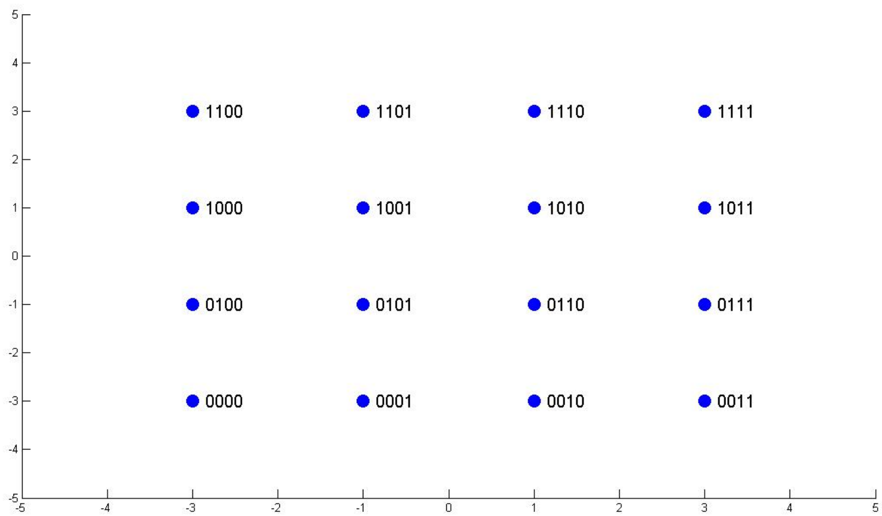

Figure 2.1: The model can produce 16-QAM

AWM is able to produce 16-QAM constellation using the weight matrix given above. It can be easily seen that this model is capable of producing most basic constellations of interest like PAM and QAM.

The equation in (2.6) can also be interpreted as linear projection of vertices of a hyper-cube using projection matrix W. As it is shown in [16] a constellation for BICM is First Order Optimal(FOO) if and only if it is a linear projection of a zero-mean hypercube.

Another way to look at coefficient matrix C beside the vertices of a hyper-cube is the following. Suppose we sort the m-digit binary expansion of integer form 0 to $2^m - 1$ in an $2^m \times m$ matrix called C_1. For m = 2 it looks like

$$
C_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{2.10}
$$

Consider ϕ to be a mapping from $\{0, 1\}$ to $\{-1, 1\}$

$$
\begin{aligned}
\phi :& \{0, 1\} \mapsto \{-1, 1\} \\
\phi(b) &= 2b - 1
\end{aligned} \tag{2.11}
$$

If ϕ is performed element-wise on C_1, the result is the matrix C defined in (2.6). For example for C_1 in (2.10)

$$
\phi(C_1) = \begin{bmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \tag{2.12}
$$

is the same as $C_{4\times2}$.

Therefore, There is a one-to-one correspondence between C and C_1 defined by one-to-one mapping ϕ. The advantage of considering C_1 is that it is equivalent to a binary labelling on the constellation. We assign a unique m-bit binary string to each constellation point. In future chapters, we will see the effect of this binary labelling on the performance of the system and discuss different binary labellings associated with a given constellation.

Lemma 1. *The columns of C are orthogonal to each other and all have the same length equal to $\sqrt{M}$. In other words*

$$
C^T \times C = \begin{bmatrix} M & 0 & \cdots & 0 \\ 0 & M & \cdots & 0 \\ & & \ddots & \vdots \\ 0 & \cdots & 0 & M \end{bmatrix} \tag{2.13}
$$

is a diagonal matrix of size $m \times m$.

Proof: It is obvious from the construction.

From now on, we just consider one-dimensional constellation produced by our model based on (2.6) for W of size $m \times 1$ denoted by weight vector $\mathbf{w}$. In one-dimensional constellation, X is an $M \times 1$ matrix and we show it by vector $\mathbf{x}$.From now on, we use these notations: X the random variable defined on the alphabet set $\mathcal{X}$, P_X the probability distribution of X, $\mathbf{x}$ the vector containing constellation points and $\mathbf{w}$ represents the weight vector.

some basic facts about the model is given in the following theorem.

Theorem 1. *1. permuting the elements of $\mathbf{w}$ does not change the resulting constellation. The constellation produced by $\mathbf{w}' = (w_{\sigma(1)}, ..., w_{\sigma(m)})$ is the same as the one generated by $\mathbf{w} = (w_1, ..., w_m)$.*

2. If we multiply any subset of elements of $\mathbf{w}$ in (-1) the constellation remains unchanged. for $I \subset \{1, 2, ..., m\}$ define $\mathbf{w}'$ by

$$w_i' = \begin{cases} -w_i & \textit{if} \quad i \in I \\ w_i & \textit{if} \quad i \in I \end{cases} \tag{2.14}$$

then

$$\mathcal{X}_{\mathbf{w}} = \mathcal{X}_{\mathbf{w}'}. \tag{2.15}$$

if σ is a permutation results in non-increasing order of elements of $\mathbf{w}$, then $\mathbf{w}'$ is a weight vector with non-increasing elements. From part(1) we can consider only weight vector with non-increasing elements :

$$w_1 \geq w_2 \geq \cdots \geq w_m \tag{2.16}$$

if $I \subset \{1, \cdots, m\}$ contains all the indices with negative elements at that position in weight vector $\mathbf{w}$, then $\mathbf{w}'$ defined in part(2) is a weight vector with non-negative elements. Without loss of generality, we can just work with weight vectors with non-negative elements

$$0 \preceq \mathbf{w} \ . \tag{2.17}$$

2.2.1 some special cases

Here, we analysis two special cases in our model.

First $w_1 = w_2 = \cdots = w_m = w$

Second

$$w_i = 2^{-i}\alpha \quad i = 1, \cdots, m \quad \text{for some} \quad \alpha > 0 \tag{2.18}$$

or equivalently

$$w_i = \frac{1}{2}\ w_{i-1} \qquad i = 2, \cdots m. \tag{2.19}$$

In the first case, $X = w \sum_{i=1}^{m} B_i$ so X has a Binomial distribution

$$X \sim B(m, \frac{1}{2}) \tag{2.20}$$

over m+1 points. Each two adjacent points come from two different realization of B_i's which only differ in one Bernouli random variable, let's say B_j.

$$w|x_i - x_{i+1}| = w|(b_1 + \cdots + 1 + \cdots + b_m) - (b_1 + \cdots - 1 + \cdots + b_m)| = 2w \ . \tag{2.21}$$

So the constellation is equi-distant in this case. If m is even, there is always one point at zero with highest probability of $\frac{\binom{n}{k}}{2^m}$.

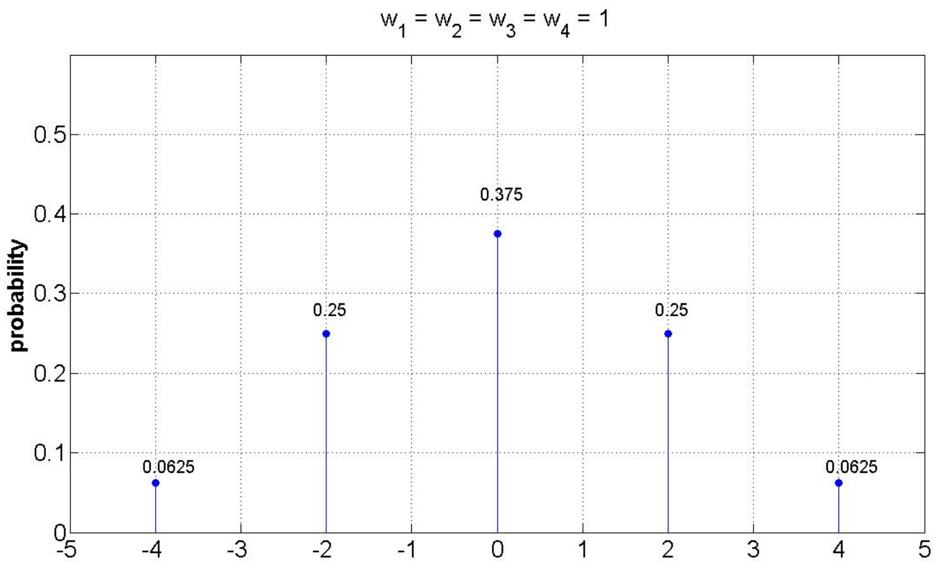

Figure 2.2: First cases: An example

In the second case, we will obtain an equi-distant equi-probable constellation. We can make any PAM constellation with equally spaced points and uniform distribution over the constellation points by choosing w_i's to satisfy (2.19).

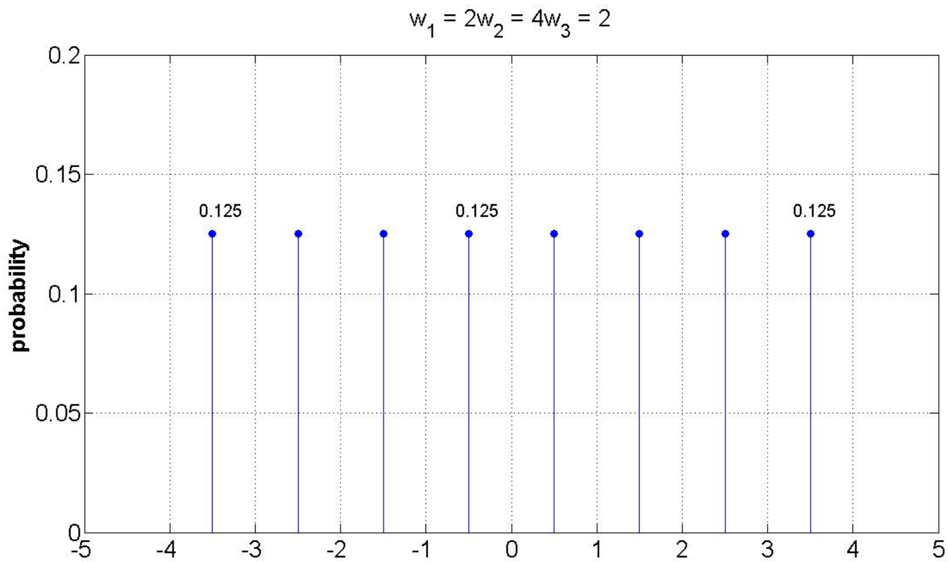

Figure 2.3: Second case: the coefficients in a dyadic form $w_i = 2^{-i}$

2.2.2 Moments

Moment generating function of a random variable X is defined to be

$$\phi_X(t) = E[e^{tX}] \tag{2.22}$$

when the expectation exists. The moments of X can be obtained from moment MGF

$$E[X^n] = \frac{d^n}{dt^n}\phi_X(t)\Big|_{t=0}. \tag{2.23}$$

Using the fact that moment generating function of sum of random variables is the product of moment generating functions of the random variables

$$E[e^{tX}] = E[e^{t(w_1 B_1)}] \cdots E[e^{t(w_m B_m)}] \tag{2.24}$$

$$E[e^{t(w_i B_i)}] = \frac{1}{2}[e^{tw_i} + e^{-tw_i}] = \cosh(tw_i) \tag{2.25}$$

so

$$\phi_X(t) = \cosh(tw_1)\cosh(tw_2)\cdots\cosh(tw_m). \tag{2.26}$$

We can simply calculate the first and second moments of the random variable defined by our model to be

$$
\begin{aligned}
E[X] = E[\sum_{i=1}^{m} w_i B_i] &= \sum_{i=1}^{m} E[w_i B_i] \\
&= \sum_{i=1}^{m} w_i E[B_i] = 0
\end{aligned}
\tag{2.27}
$$

and

$$
\begin{aligned}
E[X^2] &= E[(\sum_i w_i B_i)(\sum_j w_j B_j)] \\
&= E[\sum_{i,j} w_i w_j B_i B_j] \\
&= \sum_{i,j} w_i w_j E[B_i B_j] = \sum_i w_i^2
\end{aligned}
\tag{2.28}
$$

where (2.27) follows from $E[B_i] = 0$ and (2.28) from

$$
E[B_i B_j] = \begin{cases} E[B_i^2] = 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}
\tag{2.29}
$$

or in a more compact form

$$
E[B_i B_j] = \delta_{ij}.
\tag{2.30}
$$

From (2.28) we can see that the average power of the constellation represented by $\mathbf{w}$ is equal to the square of its Euclidean norm $\left|\left|w\right|\right|_2^2$.

In general, we have

$$
E[X^n] = \begin{cases} \sum_{i=1}^{m} w_i^n & \text{if} \quad n \ even \\ 0 & \text{if} \quad n \ odd \end{cases}
\tag{2.31}
$$

2.3 Average Power Constraint Implies Peak-Power constraint

Although average power is the only constraint we considered on the model, this constraint imposes another restriction on the placements of the points in the constellation. To obtain the maximum support of a constellation under an average power constraint P, we form the following optimization problem

$$\max_{\substack{s.t.\\ \sum_{i=1}^{m} w_i^2 \leq P \\ 0 \leq w_i}} \sum_{i=1}^{m} w_i \tag{2.32}$$

The Lagrangian of the above optimization problem [20] is

$$L(\mathbf{w}, \lambda) = \sum_i w_i - \lambda(\sum_i w_i^2 - P) \quad , \quad \lambda \geq 0 \tag{2.33}$$

To find the optimum points we try to solve

$$\nabla L = 0. \tag{2.34}$$

taking derivative with respect to each w_i

$$\frac{\partial}{\partial w_i} L = 1 - 2\lambda w_i = 0$$

$$\Rightarrow \quad w_i = \frac{1}{2\lambda} \tag{2.35}$$

so the optimum occurs at

$$w_1 = w_2 = ... = w_m = w \tag{2.36}$$

plugging (2.36) in the average power constraint

$$\sum_i w_i^2 = P \quad \Rightarrow \quad mw^2 = P$$

$$\Rightarrow \quad w_i = w = \sqrt{\frac{P}{m}} \qquad i = 1, ..., m \tag{2.37}$$

so the peak power implied by the average power constraint P is

$$\sum_{i=1}^{m} w_i = m\ w = m\sqrt{\frac{P}{m}} = \sqrt{Pm}\ . \tag{2.38}$$

Letting $m \to \infty$ for fixed P

$$\lim \sup_{m \to \infty} (X_m) = \infty$$

Chapter 3

Coded Modulation

The mutual information between two random variables indicates how much information each carries about the other one and is defined

$$I(X;Y) = D(P_{XY}||P_X\ P_Y) = E_{X,Y}\ [\log \frac{P_{Y|X}(Y|X)}{P_Y(Y)}] \tag{3.1}$$

Where expectation is taken with respect to $P_{X,Y}(X,Y)$[18]. Mutual information between two random variables can be calculated in several ways based on the definition in (3.1) which we mention some of them here.

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{3.2}$$

$$I(X;Y) = H(Y) - H(Y|X) \tag{3.3}$$

$$I(X;Y) = H(X) - H(X|Y) \tag{3.4}$$

If we face a continues random variable in calculating entropy terms, we have to use differential entropy $h(X)$ in (3.4).

According to what we have seen so far, the weight vector $\mathbf{w}$ completely characterize the constellation and its average power. Define

$$I(m, \mathbf{w}) = I(X;Y) = I(X;X+N) \tag{3.5}$$

to be the mutual information between Y, the output of the AWGN channel and the input X, which is a random variable defined over the constellation characterized by **w**. N is a zero-mean unit variance white Gaussian noise independent of signal X.

$$N \sim \mathcal{N}(0, 1). \tag{3.6}$$

Y is a mixed-gaussian random variable with at most 2^m components located at the constellation points $x_1, \cdots, x_{2^m}$. A random variable is called a mixed-gaussian random variable when its probability density function(pdf) can be written as a weighted summation of K translated Gaussian pdf.

$$P_Y(y) = \sum_{i=1}^{K} \alpha_i P_N(y - p_i) \tag{3.7}$$

for

$$\sum_{i=1}^{K} \alpha_i = 1 \qquad \alpha_i \geq 0 \qquad i = 1, \cdots, K \tag{3.8}$$

pdf for a mixed-gaussian random variable is shown in the following figure for the parameters

$$\alpha = [0.5 \quad 0.2 \quad 0.3] \tag{3.9}$$

$$\mathbf{p} = [-3 \quad 0 \quad 3]. \tag{3.10}$$

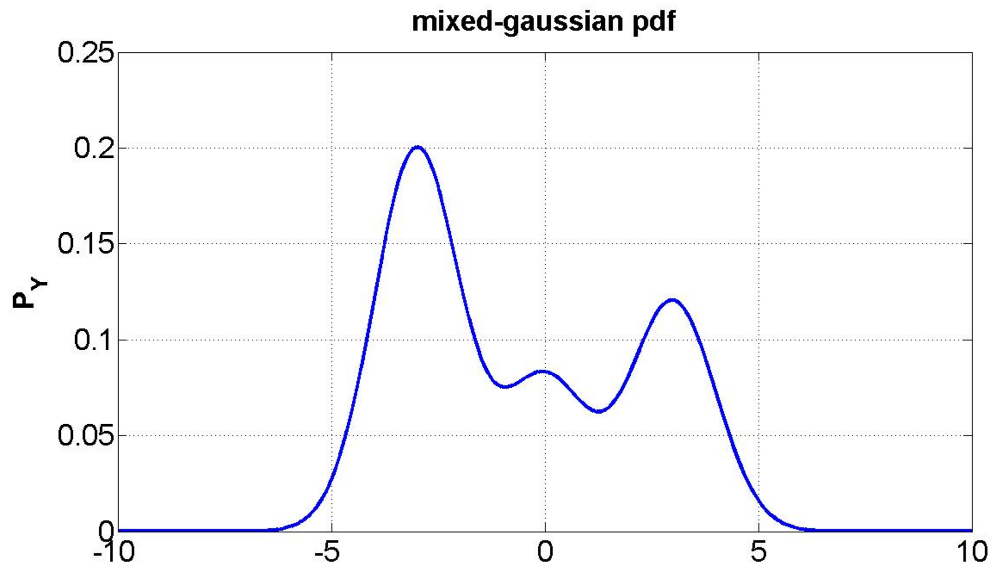

Figure 3.1: pdf for a mixed-gaussian with 3 components

Suppose X is the channel input random variable defined over $\mathcal{X} = \{x_1, \cdots, x_M\}$ with pmf $= [p_1, \cdots, p_M]$. Since the channel noise N is independent of the input X, the output pdf is

$$\begin{aligned} P_Y &= P_X * P_N \\ &= \sum_{i=1}^{M} p_i \delta(x - x_i) * P_N = \sum_{i=1}^{M} p_i P_N(x - x_i) \end{aligned} \tag{3.11}$$

and since $[p_1, \cdots, p_M]$ is a pmf

$$\sum_{i=1}^{M} p_i = 1 \qquad p_i \geq 0 \qquad i = 1, \cdots, M \tag{3.12}$$

As we see (3.11) agrees follows the condition of the definition of a mixed-gaussian distribution and hence Y is a mixed-gaussian random variable.

If input constellation is generated based on our model and assume all the points are distinct i.e. no overlapping occurs. Under this assumption, the probability distribution is uniform

$$P_X(x_i) = \frac{1}{2^m} \quad \text{for} \quad x_i \in \mathcal{X} \tag{3.13}$$

and the output pdf is

$$P_Y(y) = \frac{1}{2^m} \sum_{i=1}^{2^m} P_N(y - x_i) \tag{3.14}$$

We can calculate mutual information using

$$\begin{align} I(X;Y) &= h(Y) - h(Y|X) \tag{3.15} \\ &= h(Y) - h(X+N|X) \tag{3.16} \\ &= h(Y) - h(N) \tag{3.17} \end{align}$$

where (3.17) follows from the fact that

$$\begin{align} h(X+N|X) &= \int h(X+N|X=x) P_X(x)\ dx \tag{3.18} \\ &= h(N) \int P_X(x)\ dx = h(N). \tag{3.19} \end{align}$$

We already know the entropy of a Gaussian random variable [18]

$$h(N) = \frac{1}{2} \log\left(2\pi e \sigma^2\right). \tag{3.20}$$

We need to compute the entropy of Y to obtain the mutual information.

3.1 Coded Modulation Capacity

Given a constellation $\mathcal{X}$ the Coded Modulation capacity is defined as the maximum mutual information achieved when the constellation is used as the codebook alphabet.

From practical implementation perspective and complexity issues, one usually considers the uniform distribution:

$$P_X(x) = \frac{1}{|\mathcal{X}|} \tag{3.21}$$

$$C_{\mathcal{X}}^{CM} = I(X;Y) = E[\log \frac{P(Y|X)}{P(Y)}]. \tag{3.22}$$

In [10] authors show that we can design channel code with MLC (multi-level coding) and MSD (multi-stage decoding) that achieves coded modulation capacity by assigning appropriate rate to each level in the code. We will talk more about coded modulation capacity and how to achieve it in later sections.

3.2 Equillatice Constellation

In this section, we will look at the behaviour of the mutual information in a given snr as a function of number of input constellation points for equillatice (equi-probable equi-distant) constellation.

If we choose $x_i = i + t$ we will end up with an equi-distant constellation. Then we will find t such that $\sum_i x_i = 0$ (symmetric input constellation).

$$\sum_{i=1}^{m} x_i = \frac{m(m+1)}{2} + tm = 0 \quad \Rightarrow t = -\frac{m+1}{2} \tag{3.23}$$

$$\Rightarrow x_i = i - \frac{m+1}{2} \qquad i = 1, ..., m \tag{3.24}$$

$$\sum_i x_i^2 = \sum_i i^2 + m\frac{(m+1)^2}{4} - 2\frac{m+1}{2}\frac{m(m+1)}{2} \tag{3.25}$$

$$\sum_i x_i^2 = \sum_i i^2 - \frac{m(m+1)^2}{4} = \frac{m(m+1)(2m+1)}{6} - \frac{m(m+1)^2}{4} \tag{3.26}$$

$$= \frac{m(m+1)(m-1)}{12} \tag{3.27}$$

$$P = \frac{1}{m}\sum_i x_i^2 = \frac{1}{m}\frac{m(m+1)(m-1)}{12} = \frac{m^2-1}{12} \tag{3.28}$$

$$x_i^{'} = \frac{1}{\sqrt{P}} x_i = \frac{1}{\sqrt{P}}(i - \frac{m+1}{2}) \tag{3.29}$$

$$\Rightarrow \frac{1}{m}\sum_i x_i^{'2} = 1 \tag{3.30}$$

So $X^{'}$ is an equillatice constellation with unit average power. scaling this constellation with $\sqrt{P}$ yields an equillatice constellation with average power equal to P:

$$X = \sqrt{P}X' \quad \Rightarrow \quad \frac{1}{m}\sum_i x_i^2 = P \tag{3.31}$$

from now on, we denote the m-point equillatice constellation with $X^m = \{X_1^m, ..., X_m^m\}$. The distance between any two adjacent points in X^m is $d = \frac{\sqrt{P}}{\sqrt{a}}$
so we have: $X_i^m = d(i - \frac{m+1}{2}) \quad i = 1, ..., m$

Theorem 2. *: X^m has the following properties:*
1- It is a symmetric m-point constellation with average power equal to P:

$$E[X^m] = 0 \quad and \quad E[(X^m)^2] = P$$

2- $X_{min}^m = d\frac{1-m}{2} \quad and \quad X_{max}^m = d\frac{m-1}{2} \quad \Rightarrow X_{min}^m = -X_{max}^m$

X^m_{max} is monotone increasing in m.

3- The support of the constellation $X^m_{max} - X^m_{min}$ is increasing in m.

4- The distance between adjacent points $d = \sqrt{\frac{12P}{m^2-1}}$ is decreasing in m.

5- for odd m, there is always one point in the origin: $X^m_{\frac{m+1}{2}} = 0$

3.2.1 Optimizing the Probability Distribution

we want to optimize the rate for a 4-point equi-distant constellation over the probabilities of constellation points. we assume a symmetric constellation. so the probability mass function of the points is $P = \{p_1, p_2, p_2, p_1\}$ such that $p_i \geq 0$ and $\sum_i p_i = 1$. so

$$2p_1 + 2p_2 = 1 \qquad \text{or} \qquad p_2 = 0.5 - p_1 \tag{3.32}$$

In this case, the average power of the constellation is:

$$\sum_i p_i x_i^2 = 2p_1 x_1^2 + 2p_2 x_2^2 \tag{3.33}$$

$$= 2p_1 x_1^2 + 2(0.5 - p_1)x_2^2 \tag{3.34}$$

$$= 2p_1(x_1^2 - x_2^2) + x_2^2 \tag{3.35}$$

also since we have the equally spaced condition

$$|x_1| = |3x_2| \tag{3.36}$$

$$\Rightarrow E[X^2] = 2p_1(x_1^2 - \frac{1}{9}x_1^2) + \frac{1}{9}x_1^2 \tag{3.37}$$

$$= x_1^2(\frac{1}{9} + \frac{16}{9}p_1) \tag{3.38}$$

There are some constraints on choosing p_1, x_1 :

$$0 \leq p_1 \leq \frac{1}{2} \tag{3.39}$$

$$E[X^2] \leq P \tag{3.40}$$

Hence the optimization problem is performed on only two variables p_1 , x_1 and it can be written as:

$$\sup_{\substack{0 \le p_1 \le \frac{1}{2} \\ x_1^2(\frac{1}{9}+\frac{16}{9}p_1) \le P}} \quad I(X; X+N) \tag{3.41}$$

since $I(X; X+N)$ is a continues function and the domain is compact, the function achieves its maximum on the region defined by two constraints. Based on the result in [21] mutual information is a strictly increasing function of snr (P here). so the second inequality always holds with equality. The following optimization problem is equivalent to the original optimization problem in (3.41)

$$\sup_{\substack{0 \le p_1 \le \frac{1}{2} \\ x_1^2(\frac{1}{9}+\frac{16}{9}p_1) = P}} \quad I(X; X+N) \tag{3.42}$$

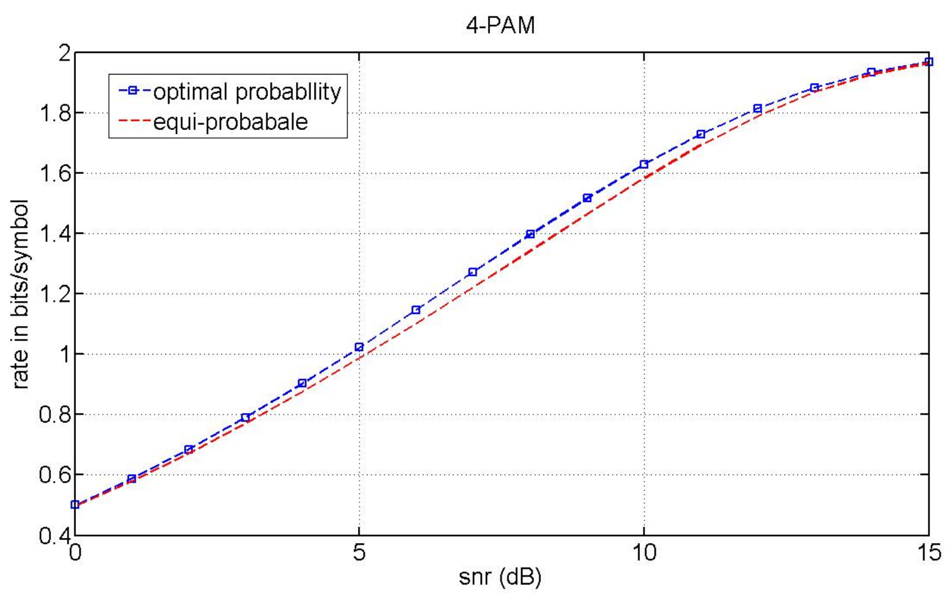

Figure 3.2: comparison between optimal probability distribution for 4-PAM constellation and uniform distribution

3.2.2 8-point constellation

If we follow the same path for an 8-point equi-distant constellation to find the optimal probability distribution, we could get the following optimization problem:

$$x_2 = 3x_1, \quad x_3 = 5x_1, \quad x_4 = 7x_1 \tag{3.43}$$

$$E[X^2] = \sum_i p_i x_i^2 \le P \quad \Rightarrow \quad 2(p_1x_1^2 + p_2x_2^2 + p_3x_3^2 + p_4x_4^2) \le P \tag{3.44}$$

or

$$2(p_1 + 9p_2 + 25p_3 + 49p_4)x_1^2 \le P \tag{3.45}$$

$$\Rightarrow x_1 \le \sqrt{\frac{P}{2(p_1 + 9p_2 + 25p_3 + 49p_4)}} \tag{3.46}$$

So the optimization problem is:

$$\sup_{\substack{0\le p_i\le \frac{1}{2} \quad i=1,..,4 \\ \sum_{i=1}^4 p_i=0.5 \\ EX^2\le P}} I(X; X+N) \tag{3.47}$$

The same reasoning as the previous section, show that the above optimization problem is equivalent to somehow easier following optimization problem:

$$\max_{\substack{p_1,p_2,p_3,p_4 \\ 0\le p_i\le \frac{1}{2} \quad i=1,..,4 \\ \sum_{i=1}^4 p_i=0.5 \\ x_1=\sqrt{\frac{P}{2(p_1+9p_2+25p_3+49p_4)}}}} I(X; X+N) \tag{3.48}$$

so for each 4-tuple (p_1, p_2, p_3, p_4) we first calculate the corresponding x_1 and hence the corresponding equidistant constellation:

$$X = [-7x_1 \quad -5x_1 \quad -3x_1 \quad -x_1 \quad x_1 \quad 3x_1 \quad 5x_1 \quad 7x_1] \tag{3.49}$$

Actually for the above optimization problem, we first reduce the domain to be the intersection of the hypercube

$$\{(p_1, p_2, p_3, p_4)| \quad 0 \leq p_i \leq 0.5 \quad i = 1, \cdots, 4\} \tag{3.50}$$

and the fist orthant of R^4 (4-dimensional Euclidean space) and again confine the search space to the points on the hyperplane $\sum_{i=1}^{4} p_i = 0.5$.

The results are shown in the following figures:

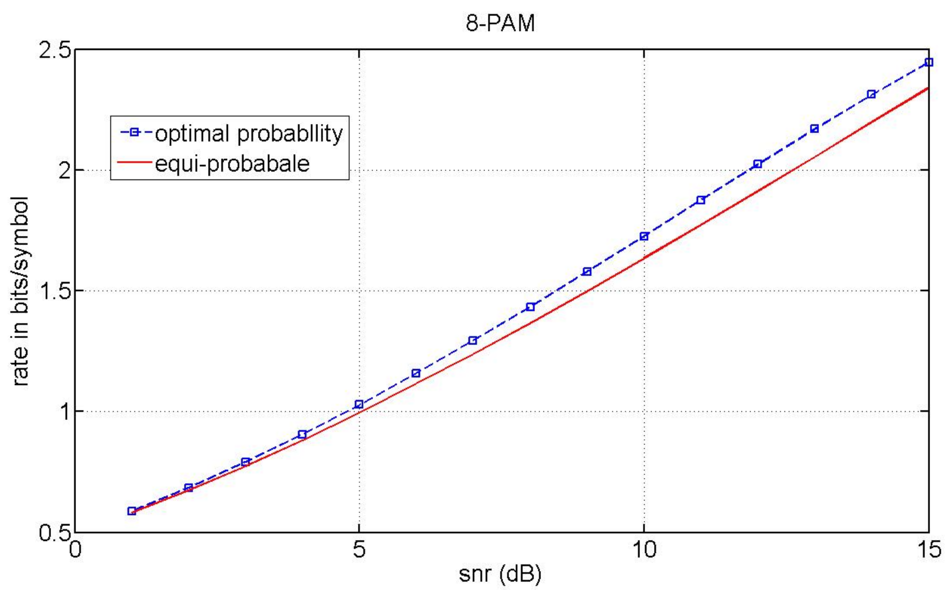

Figure 3.3: comparison between optimal probability distribution for 8-PAM constellation and uniform distribution

The rate for 8-point equillatice constellation is also shown on the same figure for comparison purposes. As we can see the rates with optimal probabilities is always greater than those for equillatice. This shaping gain is higher for medium SNR regime and decreases as snr increases. We know that for high snr regime, equillatice constellation is optimal [21], so both curves coincide with each other and as you see in the second figure, the optimal probability distribution over constellation points converges to the uniform distribution.

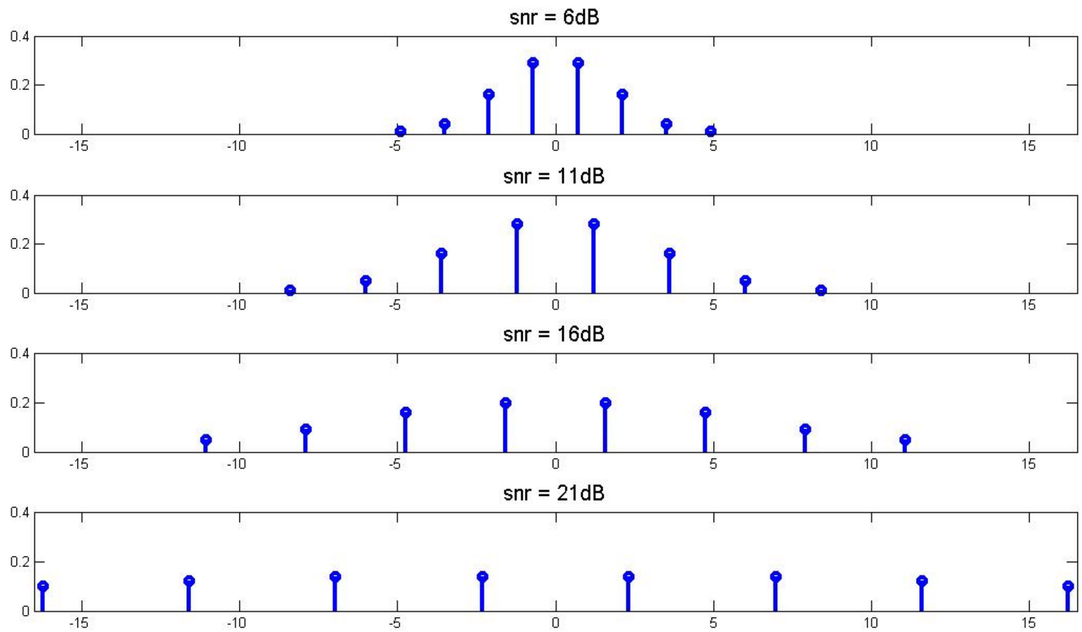

Figure 3.4: Optimal probabilities for different snr values

3.2.3 limiting behavior

In a fixed snr (fixed P) if we increase m, the input distribution looks more and more like a continues uniform distribution on $[X_{min}, X_{max}]$

$$\lim_{m\to\infty} X_{max} = \frac{\sqrt{P}}{\sqrt{a}}\frac{m-1}{2} = \frac{\sqrt{P}}{2}\frac{m-1}{\sqrt{\frac{m^2-1}{12}}} = \frac{\sqrt{12P}}{2} = \sqrt{3P} \tag{3.51}$$

$$\lim_{m\to\infty} X_{min} = -\lim_{m\to\infty} X_{max} = -\sqrt{3P} \tag{3.52}$$

In the limit the equillatice constellation converges to a uniform distribution on the interval $[-\sqrt{3P}, \sqrt{3P}]$

We can also obtain the above results using the average power constraint on $[-a, a]$:

$$E[X^2] = \int_{-a}^{a} x^2 \frac{1}{2a} dx = P \tag{3.53}$$

$$\Rightarrow \frac{1}{2a}\frac{x^3}{3}|_{-a}^{a} = \frac{1}{2a}\frac{2a^3}{3} = \frac{a^2}{3} = P. \tag{3.54}$$

$$\Rightarrow a = \sqrt{3P} \tag{3.55}$$

So X is uniformly distributed on the interval $[-\sqrt{3P}, \sqrt{3P}]$.

There is a conjecture that the mutual information is a strictly increasing function of m for any given fixed snr. if we establish the above statement to be true, we can find an upper bound for the rate of any equillatice constellation for all value of m.

$Y = X + N$. Since X, N are independent

$$\begin{aligned} P_Y &= P_X * P_N \\ P_Y(y) &= \int_{-\infty}^{\infty} P_N(t) P_X(y-t) dt \\ &= \int_{-\sqrt{3P}+y}^{\sqrt{3P}+y} P_N(t) P_X(y-t) dt \\ P_Y(y) &= \int_{-a+y}^{a+y} P_N(t) \frac{1}{2a} dt = \frac{1}{2a}[Q(y-a) - Q(y+a)] \end{aligned} \tag{3.56}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-\frac{t^2}{2}} dt \tag{3.57}$$

in this case,we follow a few steps to yield the mutual information:

$$\begin{aligned} I(X;Y) &= h(Y) - h(Y|X) \\ &= h(Y) - h(X+N|X) = h(Y) - h(N) \end{aligned} \tag{3.58}$$

Where

$$
\begin{aligned}
h(Y) &= -\int_{-\infty}^{\infty} P_Y(y) \log P_Y(y)\ dy \\
&= -\int_{-\infty}^{\infty} P_Y(y) \log \frac{1}{2a} dy - \int_{-\infty}^{\infty} P_Y(y) \log\left[Q(y-a) - Q(y+a)\right] dy \\
h(Y) &= \log(2a) - \int_{-\infty}^{\infty} P_Y(y) \log\left[Q(y-a) - Q(y+a)\right] dy
\end{aligned}
\tag{3.59}
$$

3.3 Functional Properties of Mutual Information

Mutual information in general is a functional of joint probability distribution $P_{XY}(X,Y)$ or $P_X(X)P_{Y|X}(Y|X)$. Here we present some properties of entropy and mutual information in discrete case mentioned in [21].

Mutual information is weakly continues in the input distribution P_X when the input variance is finite and the noise is additive Gaussian.

Theorem 3. *If $X_k \to X$ in distribution and sup var$X_k < \infty$ then $I(X_k, snr) \to I(X, snr)$ for any $snr \geq 0$*

Let $\mathcal{M}$ be the collection of all probability measures on (R,$\mathcal{B}$). We define $\mathcal{M}_m$ to be the collection of all the discrete random variables (constellations) which can be defined based on our model with m parameters that satisfies average power constraint.

$$
\mathcal{M}_m = \{X \in \mathcal{M}|\ X = \sum_{i=1}^{m} w_i B_i; \quad w_i \geq 0; \quad \sum_{i=1}^{m} w_i^2 \leq P\} \tag{3.60}
$$

According to [21] $\mathcal{M}_m$ is weakly compact. Since for any constellation in $\mathcal{M}_m$, $E[X^2] \leq P$ we have

$$
\sup_{X \in \mathcal{M}_m} \ E[X^2] < \infty \tag{3.61}
$$

Theorem (3) shows that $P_X \mapsto I(X, P)$ is weakly continues on $\mathcal{M}_m$, therefore it achieves its maximum on $\mathcal{M}_m$. Following Theorem (3) in [21] the optimization problem can be written as

$$C(m, P) = \max_{X \in \mathcal{M}_m} I(X, P) \tag{3.62}$$

The above discussion doesn't say anything about uniqueness of the optimal weight vectors. There might be different weight vectors that give the same maximum rate over $\mathcal{M}_m$.

Using I-MMSE relation in [21] we can show that $C(m, P)$ is strictly increasing as a function of P.

$$\frac{dI(X, snr)}{dsnr} = \frac{1}{2} MMSE(X|\sqrt{snr}X + N) \tag{3.63}$$

Note that the above relation hold when we have a constellation with unit variance $EX^2 = 1$ and Y is related to X through a AWGN channel $Y = \sqrt{snr}X + N$.

Theorem 4. *C(m,P) defined in (3.76) is a monotone increasing function of P.*

Proof: If a constellation X with average power P achieves the maximum of (3.76), and we scale it to have average power P' when $P < P'$, then using (3.63) we have $C(m, P) = I(X, P) < I(\sqrt{\frac{P'}{P}}X; Y)$. Since $I(\sqrt{\frac{P'}{P}}X; Y) \leq C(m, P')$ based on definition, we have $C_m(P) < C_m(P')$.

Lemma 2. *For every discrete real-valued random variable X*

$$H(X) = \lim_{snr \to \infty} I(X; \sqrt{snr}X + N) \tag{3.64}$$

We define the non-gaussianness of a random variable based on the Kullbak-leibler divergence between its distribution and its gaussian counterpart distribution (gaussian distribution with the same mean and variance of X). For every random variable X with finite variance $\sigma_X^2 < \infty$, its Non-Gaussianness is defined as

$$D_X = D(P_X \mid\mid \mathcal{N}(EX, \sigma_X^2)). \tag{3.65}$$

According to the formula in [21] :

$$D(Y_1||Y_2) = D(X_1 + N||X_2 + N) = I(X_1; Y_1) - I(X_2; Y_2) \tag{3.66}$$

The following theorem is useful in dealing with input distribution with finite cardinality.

Theorem 5. *If Y and Y' are output of an AWGN channel when X and X' are input random variables, respectively. The divergence between the output distributions in both cases is given by*

$$D(P_Y||P_{Y'}) = I(X'; Y') - I(X; Y) \tag{3.67}$$

Proof: see [22]

Since we know that gaussian distribution achieves capacity , if we let X' to be gaussian with the same mean and variance of X

$$\begin{aligned} D(P_Y||P_{Y'}) &= C - I(X, Y) \\ \Longrightarrow I(X; Y) &= C - D_Y \end{aligned} \tag{3.68}$$

Where D_Y is non-Gaussianness of random variable Y. If one tries to find the optimum input constellation that maximizes the mutual information achieved by it, the problem is equivalent to find the input constellation that minimizes D_Y.

3.3.1 continuity of mutual information on D(K,P)

$I(m, \mathbf{w})$ is a function of m parameters: $w_1, ..., w_m$.

$$\begin{aligned} I(\ .\ ;\ .\ ) : R^m &\mapsto R^+ \bigcup \{0\} \\ \mathbf{w} \in S(m, P) &\mapsto R^+ \bigcup \{0\} \end{aligned} \tag{3.69}$$

where

$$S(m,P) = \{\mathbf{w} = (w_1, .., w_m) \quad | \quad w_i \geq 0, \quad \sum_{i=1}^{m} w_i^2 \leq P\} \tag{3.70}$$

$$D(m,P) = \{X = \sum_{i=1}^{m} w_i B_i \quad | \quad \mathbf{w} = (w_1, .., w_m) \in S(m,P)\} \tag{3.71}$$

According to the definition, D(m,P) is the set of all constellation produced by the m-vector coming from S(m,P). We consider each constellation to be a point in a 2^m-dimensional space. So D(m,P)is a subset of R^{2^m}.

S(m,P) is the intersection of a m-dimensional sphere of radius $\sqrt{P}$ and the first orthant of m-dimensional Euclidean space. It is a closed and bounded subset of R^m, so according to the Heine-Borel theorem [20], S(m,P) is a compact set. It is also a convex set. On the other hand, D(K,P) is the set of all constellation with cardinality up to 2^m and subject to an average power constraint P which are produced based on (2.5).

D(K,P) is a subset of R^{2^m}. we want to check if it is closed and bounded. There is a linear function maps each point $\mathbf{w} \in S(m,P)$ to a point $\mathbf{x} \in D(m,P)$. D(m,P) is the image of a compact set under a continues mapping, hence it is compact.

Since each $X \in \mathcal{M}_m$ and its average power can be completely characterized by $\mathbf{w}$, we can consider mutual information as a function of $\mathbf{w} \in S(m,P)$. so

$$I(X,P) = I(m,\mathbf{w}) \tag{3.72}$$

for corresponding $\mathbf{w}$. Since $I(m,\mathbf{w})$ is a continues function on S(m,P), it achieves its maximum on S(m,P).We can form an equivalent optimization problem to the one in (3.76)

$$C(m,P) = \max_{\mathbf{w} \in S(m,P)} I(m,\mathbf{w}) \tag{3.73}$$

C(m,P) interpreted as the maximum rate (in bits per symbol) achieved when the constellation produced with m parameters is sent over an AWGN channel and the input constellation is subject to an average power constraint:

$$E[X^2] \leq P \tag{3.74}$$

The solution to the above maximization problem always exists

$$\exists \quad \mathbf{w}^* \in S(m,P) \quad s.t. \quad I(m,\mathbf{w}^*) = \max_{\mathbf{w}\in S(m,P)} I(m,\mathbf{w}). \tag{3.75}$$

If $I(m,\mathbf{w})$ was a strictly convex function in addition to continuity, the maximum point is unique,i.e. there is only one local optimum which is the same as global optimum [20]. Unfortunately as we will show it is not true in general.

There is no closed form for the mutual information function in this case. It's a highly non-linear function with a lot of local optima on S(m,P), so finding global optimum point is not easy and requires special care. Using numerical method for optimization like gradient descent is very likely to give us an sub-optimum point. Most of this gradient based algorithms are local solver which starting from an initial point, converge to one of its nearest local optima. But we are interested in characterizing $C(m,P)$ which gives us incentive to find global optima of the mutual information function on the convex compact set S(m,P). Although the domain is a convex set but unfortunately, the objective function is not a convex function, which means we can not use fast, reliable solver for convex optimization. Also we have to make sure that the output of the optimization algorithm is actually a global maximum not a local one which is not an easy task. for convex optimization problem, the local optimum is indeed the same as global optimum and we can use any local optimization solver [20].

3.4 Solving The Optimization Problem

The problem of finding the optimal weight vector that gives the maximum mutual information is a global optimization problem. The objective function might have several local optima which can be easily find by a gradient based solver. However, since we are interested in finding the maximum value the mutual information function can attain on the whole domain, the problem is a global optimization problem.

In this section, we review some common global optimization algorithms and then apply them to our problem of finding optimum weight vector. First, following [23] we show that the optimum always lies on the boundary of S(m,P).

Theorem 6. *The maximum always occurs on the boundary so we can rewrite the above optimization problem as:*

$$C(m,P) = \max_{\substack{s.t.\ w_i \geq 0 \\ \sum_{i=1}^{m} w_i^2 = P}} I(m,\mathbf{w}) \tag{3.76}$$

Proof: suppose maximum occurs at an interior point $\mathbf{w_1}$ *s.t.* $\sum_i^K w_{1i}^2 < P \quad$ *and* $I(m,\mathbf{w_1}) = C(m,P)$ *let*

$$\mathbf{w_2} = \frac{\sqrt{P}}{\sqrt{\sum_i^m w_{1i}^2}} \mathbf{w_1} = \beta\ \mathbf{w_1} \quad (\beta > 1) \tag{3.77}$$

$$\Rightarrow \sum_i^m w_{2i}^2 = P \tag{3.78}$$

if X_1 *and* X_2 *are two constellations represented by* $\mathbf{w_1}$ *and* $\mathbf{w_2}$ *respectively*

$$X_1 = \sum_{i=1}^{m} w_{1i} B_i \tag{3.79}$$

$$X_2 = \sum_{i=1}^{m} w_{2i} B_i \tag{3.80}$$

we have $X_2 = \beta X_1$.

$$Y_1 = X_1 + Z \tag{3.81}$$

$$Y_2 = X_2 + Z = \beta X_1 + Z \tag{3.82}$$

$$\tilde{Y} = \frac{1}{\beta} Y_2 + W \tag{3.83}$$

we introduced the noise term W *such that* $\tilde{Y} = Y_1$

$$\tilde{Y} = \frac{1}{\beta}[\beta X_1 + Z] + W = Y_1$$

$$\Rightarrow \quad X_1 + \frac{1}{\beta} Z + W = X_1 + Z \tag{3.84}$$

so the introduced noise term is

$$W = Z(1 - \frac{1}{\beta}) \tag{3.85}$$

for any $\beta > 1, \quad X_2 - Y_2 - \tilde{Y}$ *forms a Markov chain. Using data processing inequality, we have:*

$$\begin{aligned} I(X_2; Y_2) &= I(\beta X_1; Y_2) && (3.86)\\ &\geq I(\beta X_1; \tilde{Y}) = I(\beta X_1; Y_1) && (3.87)\\ &= I(X_1; Y_1) && (3.88)\\ \Rightarrow \quad I(X_2; Y_2) &\geq I(X_1; Y_1) && (3.89) \end{aligned}$$

The second point which is on the boundary yield higher mutual information compared to any other point strictly in the sphere.we conclude that we just need to limit our search space to the points on the boundary of S(m,P).

3.4.1 Randomized Initial Points

one way to solve the above optimization problem for finding its global maximum is to run a local solver several times, each time with a different initial points for a specific number of steps or accuracy and choose the greatest value between all the outputs. We may hope that one can overcome the problem of trapping in a local maximum, if we increase the number of times we run the solver joint with cleverly choosing initial points. we want to explore the whole search space S(m,P)which means we have to spread the initial points over the search space. There are many different algorithms regarding how to choose initial point in each run of the solver.

3.4.2 Genetic Algorithm

Genetic Algorithm (GA) is a search method that mimics the natural evolution process. It is now widely used in science and engineering for solving complicated global optimization problem. At the beginning, we randomly choose N points from the search space as our initial population which is also called first generation. We evaluate the objective function in all N points of first generation. We then discard μ portion of the points with the lowest objective function value from the first generation. We then replace the discarded population using a process called mating. The newly generated points in addition to the points remained from the first generation, all together form the second generation. We then again evaluate the objective function for the second generation and rank them based on the value of objective function at those points. We then continue the same steps of discarding and mating to form the third generation of size N. We terminate the algorithm after reaching a specific number of iteration or accuracy.

There are so many different ways to implement mating process. In general, we choose pairs of chromosome from the remaining chromosome randomly. we call each pair , parents. Each parents pair then produce two offsprings through mating process. There are also many ways to choose pairs from the remaining population as parents. One can simply choose uniformly over the remaining population. One can also assign a probability proportional to the value of the objective function at that point i.e the higher the value, the higher the chance of selecting that point as one of the parent. The simplest method for mating is single crossover. In single crossover, one point is selected for both parents strings and all the genes after that point in the parents chromosome are swapped. The resulting chromosomes are two children appear in new generation.

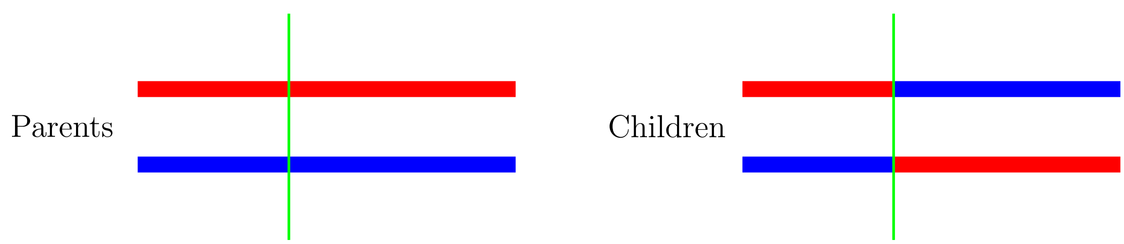

Figure 3.5: one-point crossover

Another important concept in GA is mutation. It tries to avoid GA to converge very fast to a sub-optimum solution through randomly changing the chromosomes of the remaining population in each generation.

In the following figure, you can see optimal rates for m = 3 and m = 4 achieved by solving the optimization problem using genetic algorithm. The capacity of the channel is also show for comparison purposes on the same plot. The optimal weight vector for some snr is given in the tables.

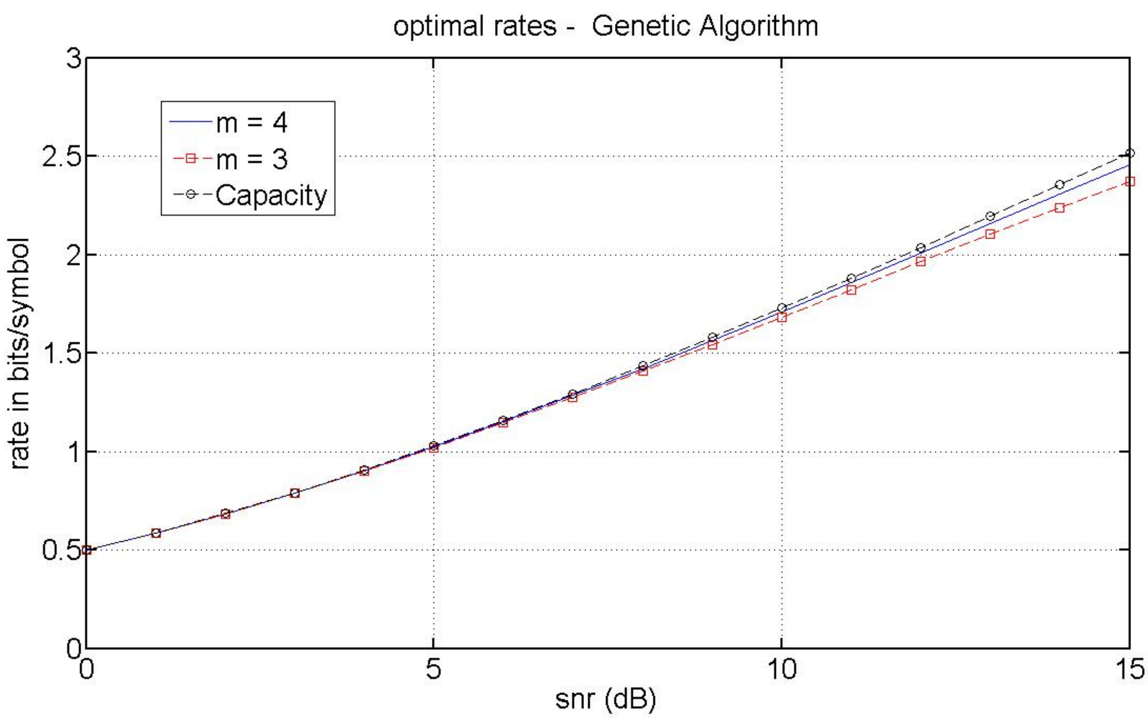

Figure 3.6: optimal rates achieved by Genetic Algorithm

	w_1	w_2	w_3	**rate**
0 dB	0.6763	0.6874	0.2647	0.4982
5 dB	0.982	1.0928	1.0018	1.0204
10 dB	1.7048	1.1794	2.388	1.5443
15 dB	3.193	4.241	1.8552	2.3706

Table 3.1: optimal weights for m = 3

	w_1	w_2	w_3	w_4	**rate**
0 dB	0.4834	0.4665	0.5843	0.4552	0.4994
5 dB	0.7932	0.8795	0.9343	0.9416	1.0249
10 dB	1.3103	1.0083	1.806	2.0012	1.7107
15 dB	3.7676	2.4502	1.5433	3.0072	2.4577

Table 3.2: optimal weights for m = 4

3.5 A New Representation of The Optimization Problem

Since the optimal point is always on the boundary (as we show in Theorem (6)) we can just limit our search space for the optimal solution that maximize rates just to be the boundary of the domain defined by the constraints of non-linear programming (3.73). For example, in 2-D we can represent each point of the boundary using polar coordinates system as

$$\begin{aligned} w_1 &= r\cos\theta \\ w_2 &= r\sin\theta. \end{aligned} \tag{3.90}$$

Since we have the following set of constraints

$$\sum_i w_i^2 = P \quad w_i \geq 0 \quad i = 1, 2 \tag{3.91}$$

we need r and θ to satisfy the following

$$r = \sqrt{P}, \qquad 0 \leq \theta \leq \frac{\pi}{2} \tag{3.92}$$

We plot the rate achieved by the points on the boundary defined by (3.92) as a function of θ for $P = 1$ and $P = 10$.

For 3-D we parametrize weight vector $\mathbf{w} = [w_1 \quad w_2 \quad w_3]$ as

$$\begin{aligned} w_1 &= r\sin\phi\cos\theta \\ w_2 &= r\sin\phi\sin\theta \\ w_3 &= r\cos\phi \end{aligned} \tag{3.93}$$

For $P = 1$ the optimal weight vector is $\mathbf{w_{opt}} = [0.5750 \quad 0.5921 \quad 0.5646]$ which occur at $[\phi \quad \theta] = [0.62 \quad 0.79]$ and the maximum achieved rate at this snr point is $r_{opt} = 0.49901345$

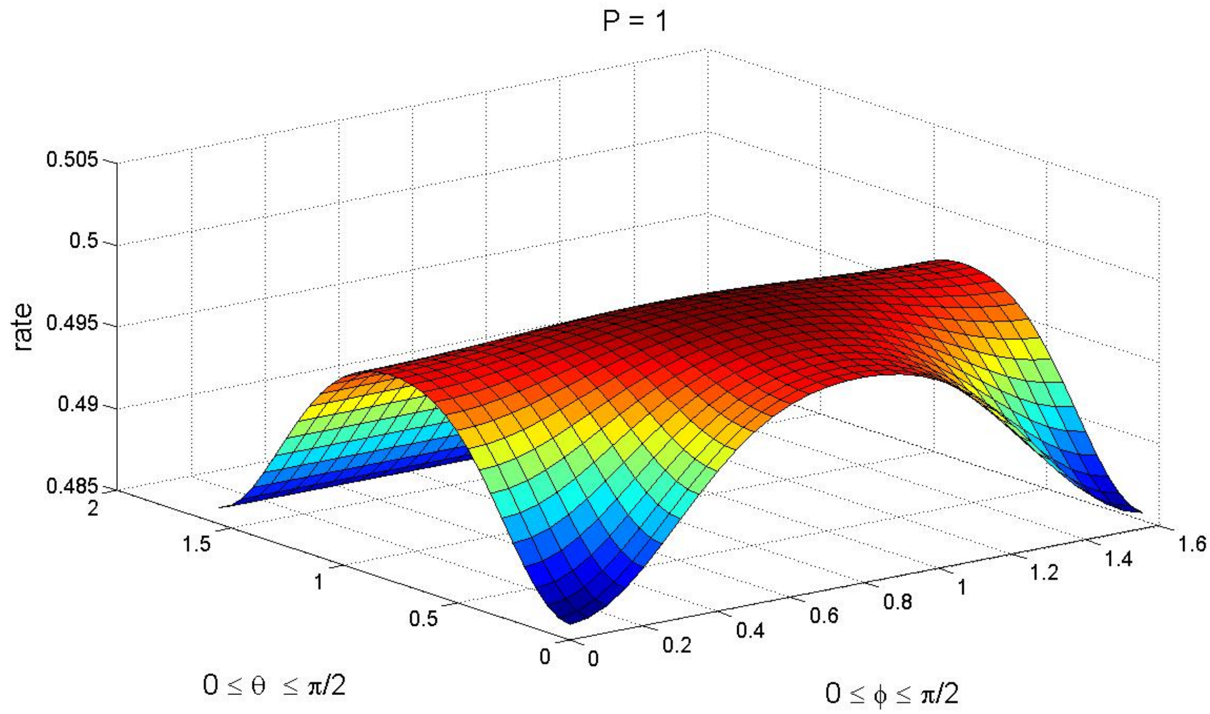

Figure 3.7: rates on the surface of $r = \sqrt{P}$ for P = 1

for $P = 10$ as you can see from the following figure there are multiple optima. One of them for example has weight vector $\mathbf{w_{opt}} = [2.346614 \quad 1.162627 \quad 1.772484]$ which occur at $[\phi \quad \theta] = [0.595 \quad 0.46]$ and the maximum achieved rate at this snr point is $r_{opt} = 1.6833267$.

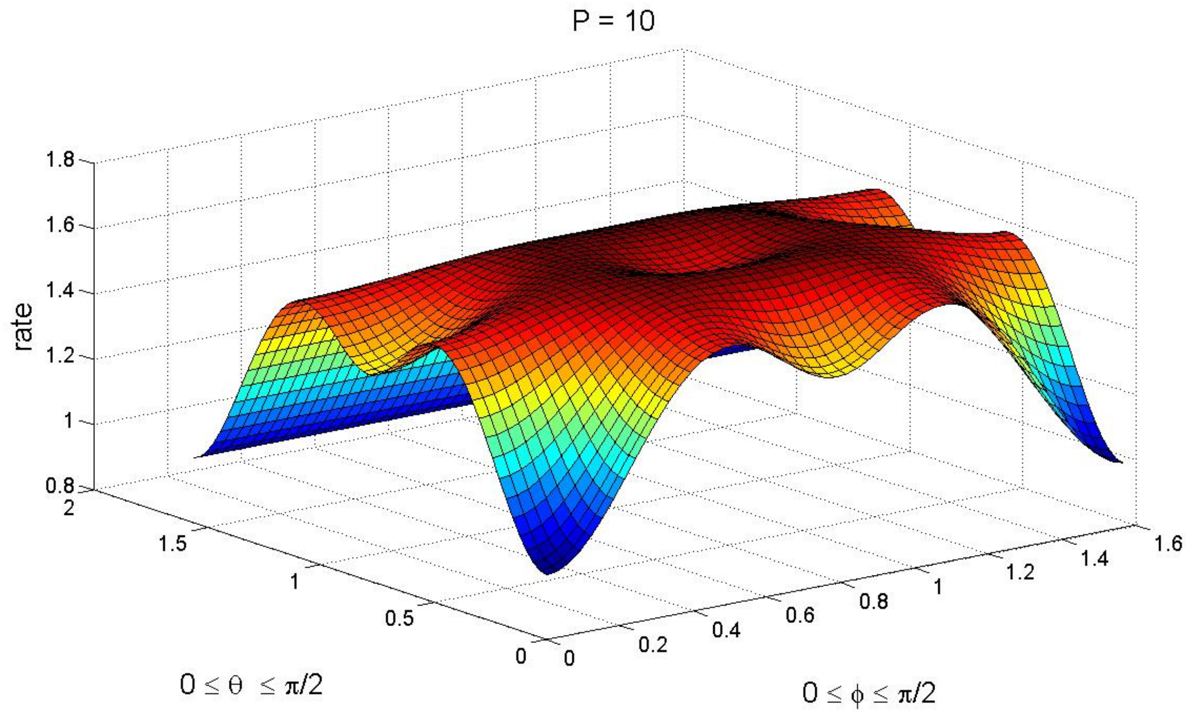

Figure 3.8: rates on the surface of $r = \sqrt{10}$

Another point which is a candidate for optimal solution is $[\phi \quad \theta] = [0.375 \quad 0.925]$ which gives the weight vector $\mathbf{w_{opt}} = [1.770912 \quad 2.349960 \quad 1.158255]$.

3.5.1 general n-dimensional case

We can define an spherical coordinate system in a n-dimensional Euclidean space R^n like what we did for 3-dimensional Euclidean space. A generalization of the previous section is as following.

There is a radial coordinate r and n-1 angular coordinates $\phi_1, ..., \phi_{n-1}$ in which ϕ_{n-1} varies between $[0 \quad 2\pi)$ radians but the first n-2 angular coordinates can only take values from $[0, \pi]$ radians. Then we can compute the weight parameters $w_1, ...w_n$ in Cartesian coordinate system from their equivalents parameters in spherical coordinates as following

$$
\begin{aligned}
w_1 =& r\cos(\phi_1) \\
w_2 =& r\sin(\phi_1)\cos(\phi_2) \\
w_3 =& r\sin(\phi_1)\sin(\phi_2)\cos(\phi_3) \\
w_{n-1} =& r\sin(\phi_1)\sin(\phi_2)...\sin(\phi_{n-2})\cos(\phi_{n-1}) \\
w_n =& r\sin(\phi_1)\sin(\phi_2)...\sin(\phi_{n-2})\sin(\phi_{n-1})
\end{aligned}
\tag{3.94}
$$

Then at each snr point we set $r = \sqrt{P}$ and solve the following optimization problem for $\phi_1, ..., \phi_{n-1}$

$$
\max_{\substack{r=\sqrt{P} \\ 0\leq\phi_{n-1}<2\pi \\ 0\leq\phi_i\leq\pi \quad i=1...n-2}} I(r, \phi_1, ..., \phi_{n-1}) \tag{3.95}
$$

and then obtaining the optimal weight vectors is just an easy transformation from spherical coordinates back to Cartesian coordinates using (3.94). After this step, we just need to use a linear transformation the one like in (2.6) to get the optimal 2^n-point input constellation.

3.6 Derivative Expressions

In this section we derives some useful expressions for derivative of mutual information with respect to weight parameters around zero which facilitate analysis in later sections. We first start with a simple example and then generalize the result in a theorem.

Suppose $\mathbf{w} = (w_1, w_2)$ with $w_1 > w_2 \geq 0$, so the input constellation is

$$
X = [-w_1 - w_2 \quad -w_1 + w_2 \quad w_1 - w_2 \quad w_1 + w_2] \tag{3.96}
$$

The channel output pdf is

$$f_Y(y) = \frac{1}{4\sqrt{2\pi}}[e^{-\frac{(y+w_1+w_2)^2}{2}} + e^{-\frac{(y+w_1-w_2)^2}{2}} + e^{-\frac{(y-w_1+w_2)^2}{2}} + e^{-\frac{(y-w_1-w_2)^2}{2}}]$$

$$\begin{aligned} f_Y(y) &= \frac{1}{4\sqrt{2\pi}}e^{-\frac{y^2+(w_1+w_2)^2}{2}}[e^{-y(w_1+w_2)} + e^{y(w_1+w_2)}] \\ &+ \frac{1}{4\sqrt{2\pi}}e^{-\frac{y^2+(w_1-w_2)^2}{2}}[e^{-y(w_1-w_2)} + e^{y(w_1-w_2)}] \end{aligned} \tag{3.97}$$

we can rewrite the pdf as

$$f_Y(y) = \frac{2}{4\sqrt{2\pi}}[e^{-\frac{y^2+(w_1+w_2)^2}{2}}\cosh y(w_1+w_2) + e^{-\frac{y^2+(w_1-w_2)^2}{2}}\cosh y(w_1-w_2)] \tag{3.98}$$

$$I(X;Y) = h(Y) - h(Z) \tag{3.99}$$

Taking derivative with respect to w_2 yields

$$\frac{\partial I}{\partial w_2} = \frac{\partial}{\partial w_2}h(Y) - \frac{\partial}{\partial w_2}h(Z) = \frac{\partial}{\partial w_2}h(Y) \tag{3.100}$$

$$= -\int \frac{\partial f_Y}{\partial w_2}(1 + \log f_Y)\ dy \tag{3.101}$$

But

$$\int \frac{\partial f_Y}{\partial w_2}\ dy = \frac{\partial}{\partial w_2}\int f_Y\ dy = \frac{\partial}{\partial w_2}1 = 0 \tag{3.102}$$

So

$$\begin{aligned} \frac{\partial f_Y}{\partial w_2} &= \frac{-2(w_1+w_2)}{2\sqrt{2\pi}}e^{-\frac{y^2+(w_1+w_2)^2}{2}}\cosh y(w_1+w_2) \\ &+ \frac{y\sinh(y(w_1+w_2))}{2\sqrt{2\pi}}e^{-\frac{y^2+(w_1+w_2)^2}{2}} \\ &+ \frac{2(w_1-w_2)}{2\sqrt{2\pi}}e^{-\frac{y^2+(w_1-w_2)^2}{2}}\cosh y(w_1-w_2) \\ &- \frac{y\sinh(y(w_1-w_2))}{2\sqrt{2\pi}}e^{-\frac{y^2+(w_1-w_2)^2}{2}} \end{aligned} \tag{3.103}$$

Evaluating partial derivative with respect to w_2 at $w_2 = 0$

$$\begin{aligned}
\left.\frac{\partial f_Y}{\partial w_2}\right|_{w_2=0} &= \frac{1}{2\sqrt{2\pi}} e^{-\frac{(y^2+w_1^2)}{2}} [-2w_1 \cosh yw_1 \\
&+ y \sinh yw_1 + 2w_1 \cosh yw_1 - y \sinh yw_1] \\
&= 0 \qquad (3.104)
\end{aligned}$$

Therefore

$$\left.\frac{\partial I}{\partial w_2}\right|_{\substack{w_2=0 \\ w_1>0}} = 0 \qquad (3.105)$$

We now state the generalization of the above example in the following theorem. Based on the above observations we have the following theorem:

Theorem 7. *For any k, the derivative of mutual information with respect to w_k at $w_k = 0$ is zero:*

$$\left.\frac{\partial I}{\partial w_k}\right|_{w_k=0} = 0 \qquad (3.106)$$

Proof:

$$f_Y(y) = \frac{1}{2^k} \sum_{i=1}^{2^k} f_N(y - \mu_i) \qquad (3.107)$$

$$\frac{\partial}{\partial w_k} f_Y = \frac{1}{2^k} \sum_{i=1}^{2^k} \frac{\partial}{\partial \mu_i} f_N(y - \mu_i) \frac{\partial}{\partial w_k} \mu_i \qquad (3.108)$$

$$= \frac{1}{2^k} \sum_{i=1}^{2^k} (-1)^{k+1} \frac{(y - \mu_i)}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\mu_i)^2} \qquad (3.109)$$

if $w_k = 0$ we have 2^{k-1} distinct points: $\mu_1 = \mu_2, \quad \mu_3 = \mu_4 \quad ...., \quad \mu_{2^k-1} = \mu_{2^k}$

$$\begin{aligned}
\Rightarrow \frac{(y - \mu_{2i-1})}{\sqrt{2\pi}} & e^{-\frac{1}{2}(y-\mu_{2i-1})^2} \\
&= \frac{(y - \mu_{2i})}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\mu_{2i})^2} \qquad i = 1, ..., 2^{k-1} \qquad (3.110)
\end{aligned}$$

So

$$\frac{\partial}{\partial w_k} f_Y|_{w_k=0} = 0 \tag{3.111}$$

$$\Rightarrow \frac{\partial I}{\partial w_k}|_{w_k=0} = 0 \tag{3.112}$$

3.7 Finding Optimum Parameters Using Moment Matching

Since the Gaussian distribution is the optimal input distribution for AWGN channel that maximizes the achievable rate we are trying to find weight parameters w_i in a way that the output distribution looks like Gaussian or in other words the difference between the output distribution and its Gaussian counterpart is as small as possible. We can consider different metrics as measuring the deviation form Gaussian. Here following the work in [22] we are trying to match as many moments as possible with moments of Gaussian distribution. The moments of the distribution in AWM are simply characterized in (2.31). X and N are independent,hence

$$E[Y^k] = E[X^k] + E[N^k]. \tag{3.113}$$

Since we have m parameters, we can find parameter vector by solving the following set of equations for the first m non-zero moments. Since odd moments of a Gaussian distribution are zero similar to AWM, we can match the first 2m moments of both distributions and form the following system of equations:

$$\begin{cases} w_1^2 + .. + w_m^2 + E[N^2] = E[X_g^2] \\ w_1^4 + .. + w_m^4 + E[N^4] = E[X_g^4] \\ . \\ . \\ w_1^{2m} + ... + w_m^{2m} + E[N^{2m}] = E[X_g^{2m}] \end{cases} \tag{3.114}$$

Where X_g is a Gaussian distribution with the same mean and variance as the output distribution Y

$$E[X_g^k] = \begin{cases} \sigma^p (k-1)!! & \text{if } k \quad \textit{even} \\ 0 & \text{if } k \quad \textit{odd} \end{cases} \tag{3.115}$$

where n!! denotes the double factorial, which is the product of all odd number from 1 to n. If we simplify (3.114) for case m = 3 when the average power constraint is P we would have:

$$\begin{cases} w_1^2 + w_2^2 + w_3^2 + 1 & = \sigma^2 = P + 1 \\ w_1^4 + w_2^4 + w_3^4 + 3 & = 3\sigma^4 = 3(P+1)^2 \\ w_1^6 + w_2^6 + w_3^6 + 15 & = 15\sigma^6 = 15(P+1)^3 \end{cases} \tag{3.116}$$

unfortunately the above system of non-linear equations does not have any real solution.

We take another approach to find a weight vector using moment matching strategy. We select $\mathbf{w}$ to satisfy the first equation in (3.114) and try to minimize the difference between LHS and RHS of other equations.

$$\min_{w_1^2+w_2^2+w_3^2=P} |w_1^4 + w_2^4 + w_3^4 + 3 - 3(P+1)^2| \tag{3.117}$$

since

$$w_1^4 + w_2^4 + w_3^4 = (w_1^2 + w_2^2 + w_3^2)^2 \tag{3.118}$$

$$- 2w_1^2w_2^2 - 2w_1^2w_3^2 - 2w_2^2w_3^2 \tag{3.119}$$

$$= P^2 - 2w_1^2w_2^2 - 2w_1^2w_3^2 - 2w_2^2w_3^2 \tag{3.120}$$

$$|w_1^4 + w_2^4 + w_3^4 + 3 - 3(P+1)^2| \tag{3.121}$$

$$= |P^2 - 2w_1^2w_2^2 - 2w_1^2w_3^2 - 2w_2^2w_3^2 + 3 - 3P^2 - 6P - 3| \tag{3.122}$$

$$= | - 2P^2 - 2w_1^2w_2^2 - 2w_1^2w_3^2 - 2w_2^2w_3^2 - 6P| \tag{3.123}$$

$$= 2P^2 + 6P + 2w_1^2w_2^2 + 2w_1^2w_3^2 + 2w_2^2w_3^2 \tag{3.124}$$

Since P is fixed, the optimization problem in (3.117) is equivalent to

$$\min_{w_1^2+w_2^2+w_3^2=P} w_1^2w_2^2 + w_1^2w_3^2 + w_2^2w_3^2. \tag{3.125}$$

We solve (3.125) by Lagrange multipliers method. The Lagrangian of the above optimization problem is

$$L(\mathbf{w}, \lambda) = w_1^2w_2^2 + w_1^2w_3^2 + w_2^2w_3^2 - \lambda(w_1^2 + w_2^2 + w_3^2 - P) \tag{3.126}$$

$$\frac{\partial L}{\partial w_1} = 2w_1w_2^2 + 2w_1w_3^2 - 2\lambda w_1 = 0 \tag{3.127}$$
$$\frac{\partial L}{\partial w_2} = 2w_2w_1^2 + 2w_2w_3^2 - 2\lambda w_2 = 0 \tag{3.128}$$
$$\frac{\partial L}{\partial w_3} = 2w_3w_1^2 + 2w_3w_2^2 - 2\lambda w_3 = 0 \tag{3.129}$$

Since we assume all weight parameters are non-zero (positive)

$$w_2^2 + w_3^2 = \lambda \tag{3.130}$$
$$w_2^1 + w_3^2 = \lambda \tag{3.131}$$
$$w_1^2 + w_2^2 = \lambda \tag{3.132}$$

Therefore we have

$$w_1^2 = w_2^2 = w_3^2$$

or

$$w_2 = w_2 = w_3. \tag{3.133}$$

The result holds for general case. If we have m parameters, the weight vector that minimizes the difference between the fourth moments is

$$w_1 = w_2 = \cdots = w_m \tag{3.134}$$

3.7.1 Cumulant Matching

The cumulants of a random variable X, denoted as k_m, are the coefficients of the following series expansion for the logarithm of characteristic function

$$\log \phi(t) = \sum_{m=0}^{\infty} k_m \frac{(it)^m}{m!} \tag{3.135}$$

where $\phi(t)$ is the characteristic function of random variable X, defined as

$$\phi(t) = E[e^{itX}] = \int_{-\infty}^{\infty} f(x)\ e^{itx} dx \tag{3.136}$$

There are simple expressions for the first few cumulants based on the moments of the random variable.

$$\begin{aligned} k_0 &= 1 \\ k_1 &= \mu_1 \\ k_2 &= \mu_2 - \mu_1^2 \\ k_3 &= 2\mu_1^3 - 3\mu_1\mu_2 + \mu_3 \\ k_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4 \end{aligned} \tag{3.137}$$

using (3.137) we can evaluate the first few cumulants for our model with weight vector $\mathbf{w} = (w_1, .., w_m)$

$$\begin{aligned} k_0 &= 1 \\ k_1 &= \mu_1 = 0 \\ k_2 &= \mu_2 - \mu_1^2 = \sum_{i=1}^{m} w_i^2 \\ k_3 &= \mu_3 = \sum_{i=1}^{m} w_i^3 \\ k_4 &= \mu_4 - 3\mu_2^2 = \sum_{i=1}^{m} w_i^4 - 3\sum_{i=1}^{m} w_i^2 \end{aligned} \tag{3.138}$$

For a Gaussian random variable only the first two cumulants k_1, k_2 are non-zero [24]. For a standard normal noise N

$$k_1(N) = 0 \tag{3.139}$$

$$k_2(N) = 1 \tag{3.140}$$

$$k_i(N) = 0 \qquad i = 3, 4\cdots \tag{3.141}$$

So

$$k_1(Y) = k_1(X) \tag{3.142}$$

$$k_2(Y) = k_2(X) + 1 \tag{3.143}$$

$$k_i(Y) = k_i(X) \qquad i = 3, 4\cdots \tag{3.144}$$

Since the cumulants of a Gaussian distribution with the same mean and variance of Y, $\mathcal{N} \sim (E[Y], \sigma_Y^2)$ are zero for order 3 and above, we are trying to minimize cumulants of Y of order 3 and above. We can match the first two cumulants easily. If there are multiple minimizer for each cumulant say k_p, we choose the one that minimizes the absolute value of the next cumulant k_{p+1}.

$$\min_{s.t. \sum_{i=1}^{m} w_i^2 = P} \qquad |k_3| = |\sum_{i=1}^{m} w_i^3| = \sum_{i=1}^{m} w_i^3 \tag{3.145}$$

The Lagrangian of the above optimization problem is

$$L(\mathbf{w}, \lambda) = \sum_{i=1}^{m} w_i^3 - \lambda(\sum_{i=1}^{m} w_i^2 - P) \tag{3.146}$$

$$\frac{\partial L}{\partial w_i} = 3w_i^2 - 2\lambda w_i \qquad i = 1, \cdots m \tag{3.147}$$

$$w_i = \frac{2\lambda}{3} \qquad i = 1, \cdots m \tag{3.148}$$

$$w_1 = w_2 = \cdots = w_m = \frac{2\lambda}{3} \tag{3.149}$$

Which yield the same result using moment matching approach.

3.7.2 Using Taylor Expansion

Mutual Information in general is a functional of joint probability distribution of input and output. But if the channel ($P_{Y|X}(Y|X)$) is fixed and we just scale the constellation to satisfy the average power constraint at the input, we can consider the mutual information to be a function of snr. We can write it as a Taylor series expansion around zero snr. By using Taylor expansion of mutual information $I(X, snr)$ around zero snr given in [22] for constellation with $E[X] = 0$ and $E[X^2] = 1$

$$I(X, snr) = \frac{\log e}{2}[snr - \frac{1}{2}snr^2 + (2 - (EX^3)^2)\frac{snr^3}{6}$$
$$+(-15 + 12(EX^3)^2 + 6EX^4 - (EX^4)^2)\frac{snr^4}{24}] + O(snr^5) \tag{3.150}$$

Trying to maximize the mutual information in low snr regime is equivalent to maximizing the coefficients of the low terms of expansion in (3.150). Since the coefficients up to order 3 are already matched we form an optimization problem to maximize the coefficient of $\frac{snr^4}{24}$ term.

$$\max_{s.t.\ E[X^2]=1} \quad 6E[X^4] - (E[X^4])^2 \tag{3.151}$$

$$\sum_{i=1}^{m} w_i^4 = (\sum_{i=1}^{m} w_i^2)^2 - 2(w_1^2 w_2^2 + w_1^2 w_3^2 + \cdots + w_{m-1}^2 w_m^2) \tag{3.152}$$

so

$$E[X^4] = 1 - 2(w_1^2 w_2^2 + w_1^2 w_3^2 + \cdots + w_{m-1}^2 w_m^2) \tag{3.153}$$

let

$$T = (w_1^2 w_2^2 + w_1^2 w_3^2 + \cdots + w_{m-1}^2 w_m^2) \tag{3.154}$$

$$\max_{E[X^2]=1} 6(1-2T)-(1-2T)^2 \tag{3.155}$$

$$\max_{E[X^2]=1} 5-4T^2-8T \tag{3.156}$$

$$\min_{E[X^2]=1} 4T^2+8T \tag{3.157}$$

Lagrangian

$$L(\mathbf{w},\lambda)=4T^2+8T-\lambda(E[X^2]-1) \tag{3.158}$$

$$\frac{\partial L}{\partial w_i}=8T\frac{\partial T}{\partial w_i}+8\frac{\partial T}{\partial w_i}-2\lambda w_i \tag{3.159}$$

$$\frac{\partial T}{\partial w_i}=2w_i(\sum_{\substack{j=1\\j\neq i}}^{m} w_j^2) \tag{3.160}$$

$$\frac{\partial L}{\partial w_i}=0 \tag{3.161}$$

$$\Rightarrow (T+1)(\sum_{\substack{j=1\\j\neq i}}^{m} w_j^2)=\mu \qquad i=1,\cdots,m \tag{3.162}$$

The solution to (3.162) is

$$w_1=w_2=\cdots=w_m \tag{3.163}$$

3.8 Properties of Coded Modulation Capacity

Suppose a communication system consists of a binary channel encoder followed by a modulator. We take each m-bit string of encoder output and map it to a constellation point through a labelling function $\mu : \{0,1\}^m \mapsto \mathcal{X}$. We can indicate each m-bit string of the encoder output by $B_1 B_2 \cdots B_m$. Since μ is a one-to-one mapping, the mutual information between the channel input X and the output Y is equal to the mutual information between the encoder output $B_1 B_2 \cdots B_m$ and the channel output Y

$$I(X;Y) = I(B_1 B_2 \cdots B_m; Y) \tag{3.164}$$

Expanding the mutual information in (3.164) using chain rule [18] gives

$$I(B_1 B_2 \cdots B_m; Y) = I(B_1; Y) + I(B_2; Y|B_1) + \cdots I(B_m; Y|B_1 B_2 \cdots B_{m-1}). \tag{3.165}$$

Expansion in (3.165) is the main essence of MLC-MSD (Multi-Level Coding Multi-Stage Decoding) first proposed by Imai and Hirakawa [10] to achieve coded modulation capacity. We can design m binary codes each with appropriate rate associated with $I(B_k; Y|B_1 \cdots B_{k-1})$ in a multilayer fashion. Decoding is performed in a sequential manner. The first bit is decoded first while the other bits are considered noise in decoding B_1. In decoding B_k the first k-1 bits $B_1 \cdots B_{k-1}$ are know and presented to the decoder. Decoder is provided with the knowledge of already decoded bits in each stage.

In this method, if each code rate is adjusted to the corresponding term in the expansion of (3.165)

$$r_k = I(B_k; Y|B_1 \cdots B_{k-1}) \tag{3.166}$$

Then

$$R = \sum_{k=1}^{m} r_k. \tag{3.167}$$

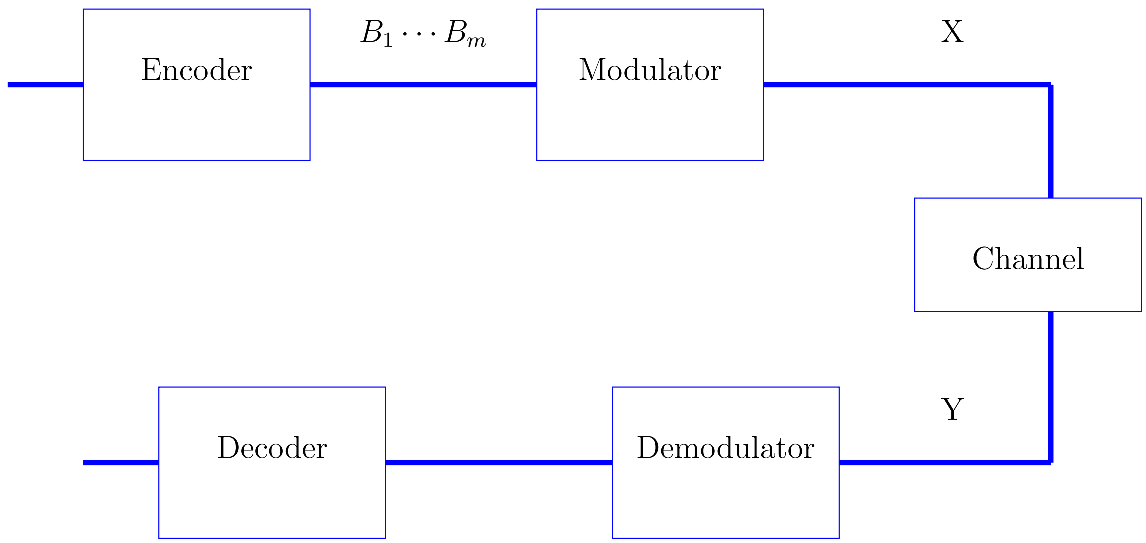

Figure 3.9: Coded Modulation Block Diagram

There are some questions might arise regarding Coded Modulation and MLC-MSD implementation of it. Chain rule in (3.165) can be written in many different ways where each expansion can be interpreted as a decoding order of bits in the decoder. For example, consider two different ways of expanding $I(B_1B_2;Y)$

$$I(B_1B_2;Y) = I(B_1;Y) + I(B_2;Y|B_1) \tag{3.168}$$

$$I(B_1B_2;Y) = I(B_2;Y) + I(B_1;Y|B_2) \tag{3.169}$$

The first one corresponds to a MLC scheme that decodes B_1 first and then passes it to the next stage to decode B_2. The rate for the code that encodes B_1 is $r_1 = I(B_1;Y)$ and the rate of the code that encodes B_2 is $r_2 = I(B_2;Y|B_1)$ for this decoding order in the decoder.

However, the second one corresponds to a MLC scheme that decodes B_2 first and then passes it to the next stage to decode B_1. The rate for the code that encodes B_1 is $r'_1 = I(B_1;Y|B_2)$ and the rate of the code that encodes B_2 must be set to $r'_2 = I(B_2;Y)$ for this decoding order if we want to achieve reliable communication.

In general these quantities are not equal

$$r'_1 \neq r_1 \tag{3.170}$$

$$r'_2 \neq r_2 \tag{3.171}$$

i.e. rate associated to each level is different for each decoding order.

Does that affect the overall performance of the system? Chain rule answers this question. Although, the rate associated to each level might be different in each order of decoding, but since

$$\sum_{i=1}^{m} r_i = \sum_{i=1}^{m} r'_i \tag{3.172}$$

the overall rate associated to Coded Modulation scheme is independent of the order in which bits are decoded in the decoder as long as the rates for each level chosen correctly.

In the following example we will see this phenomenon in more details. Suppose we use a 4-PAM constellation with a specific labelling, say Natural labelling to see the effect of order of decoding.

$$\mathcal{X} = \{-3 \quad -1 \quad +1 \quad +3\} \tag{3.173}$$

with labelling

$$\{00 \quad 01 \quad 10 \quad 11\}. \tag{3.174}$$

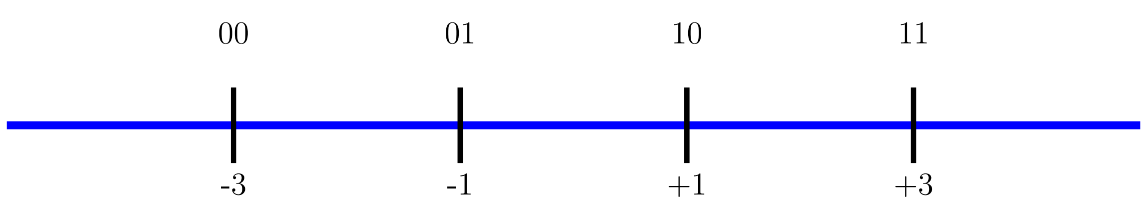

Figure 3.10: 4-PAM with Natural labelling

Calculating both terms in (3.168) leads to the followings

$$\begin{aligned} I(Y;B_1) &= h(Y) - h(Y|B_1) \\ &= h(Y) - [h(Y|B_1=0)P_{B_1}(0) + h(Y|B_1=1)P_{B_1}(1)] \end{aligned} \tag{3.175}$$

$$\begin{aligned} I(B_2;Y|B_1) &= h(Y|B_1) - h(Y|B_1B_2) \\ &= h(Y|B_1=0)P_{B_1}(0) + h(Y|B_1=1)P_{B_1}(1) \\ &- [h(Y|B_1B_2=00)P_{B_1B_2}(0,0) \\ &+ h(Y|B_1B_1=01)P_{B_1B_2}(0,1) \\ &+ h(Y|B_1B_2=10)P_{B_1B_2}(1,0) \\ &+ h(Y|B_1B_2=11)P_{B_1B_2}(1,1)] \end{aligned} \tag{3.176}$$

for the first term:

$$\begin{aligned} &P_{B_1}(0)[h(Y|B_1=0) - (P_{B_2}(0)h(Y|B_1B_2=00) + P_{B_2}(1)h(Y|B_1B_2=01))] \\ &= P_{B_1}(0)[h(Y|B_1=0) - h(Y|B_1=0,B_2)] = P_{B_1}(0)I(B_2;Y|B_1=0) \end{aligned} \tag{3.177}$$

and for the second term:

$$\begin{aligned} &P_{B_1}(1)[h(Y|B_1=1) - (P_{B_2}(0)h(Y|B_1B_2=10) + P_{B_2}(1)h(Y|B_1B_2=11))] \\ &= P_{B_1}(1)[h(Y|B_1=1) - h(Y|B_1=1,B_2)] = P_{B_1}(1)I(B_2;Y|B_1=1) \end{aligned} \tag{3.178}$$

so equation (3.176) can be rewritten as:

$$I(B_2;Y|B_1) = P_{B_1}(0)I(B_2;Y|B_1=0) + P_{B_1}(1)I(B_2;Y|B_1=1) \tag{3.179}$$

In this case Y is a mixed-Gaussian random variable with four components at constellation points. $h(Y|B_1=0)$ is the entropy of a mixed-Gaussian random variable with two components at -3, -1 and $h(Y|B_1=1)$ is the entropy of a mixed-Gaussian random variable with two components at +1, +3. $I(B_2;Y|B_1=0)$ is the mutual information obtained by transmitting a two-point constellation over an AWGN channel with points at -3, -1 and

$I(B_2; Y|B_1 = 1)$ is the mutual information obtained by transmitting a two-point constellation over an AWGN channel with points at $+1$, $+3$. The rates obtained by this particular decoding order is shown in the following figure:

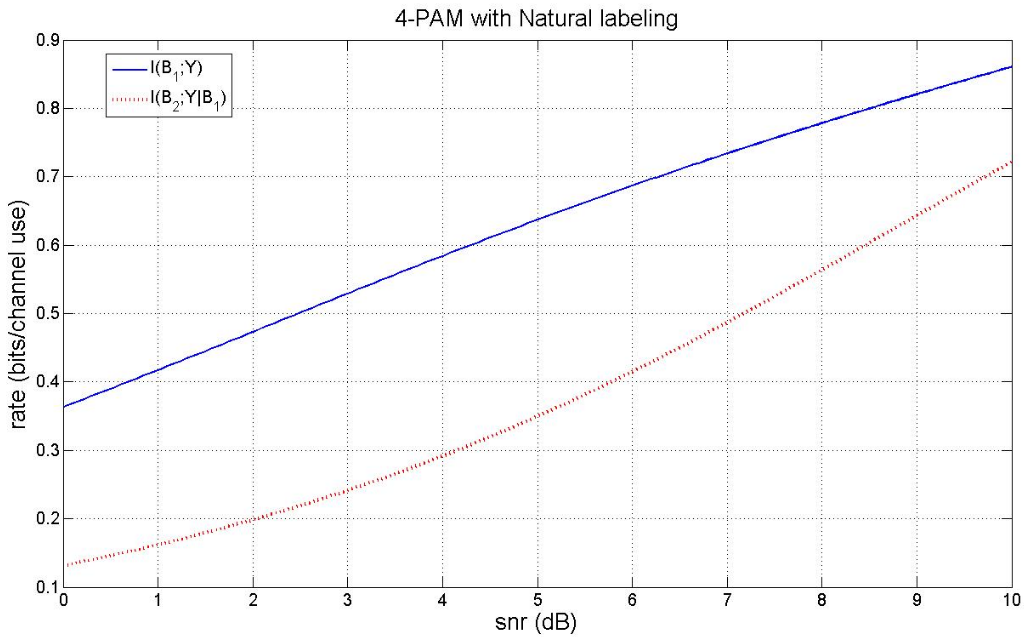

Figure 3.11: first order of decoding given in (3.168)

If we choose another decoding order in this example, we would have:

$$I(B_1 B_2; Y) = I(B_2; Y) + I(B_1; Y|B_2) \tag{3.180}$$

$$\begin{aligned} I(Y; B_2) &= h(Y) - h(Y|B_2) \\ &= h(Y) - [h(Y|B_2 = 0)P_{B_2}(0) + h(Y|B_2 = 1)P_{B_2}(1)] \end{aligned} \tag{3.181}$$

$$\begin{aligned} I(B_1; Y|B_2) &= I(B_1; Y|B_2 = 0)P_{B_2}(0) + I(B_1; Y|B_2 = 1)P_{B_2}(1) \\ &= \frac{1}{2}[I(B_1; Y|B_2 = 0) + I(B_1; Y|B_2 = 1)] \end{aligned} \tag{3.182}$$

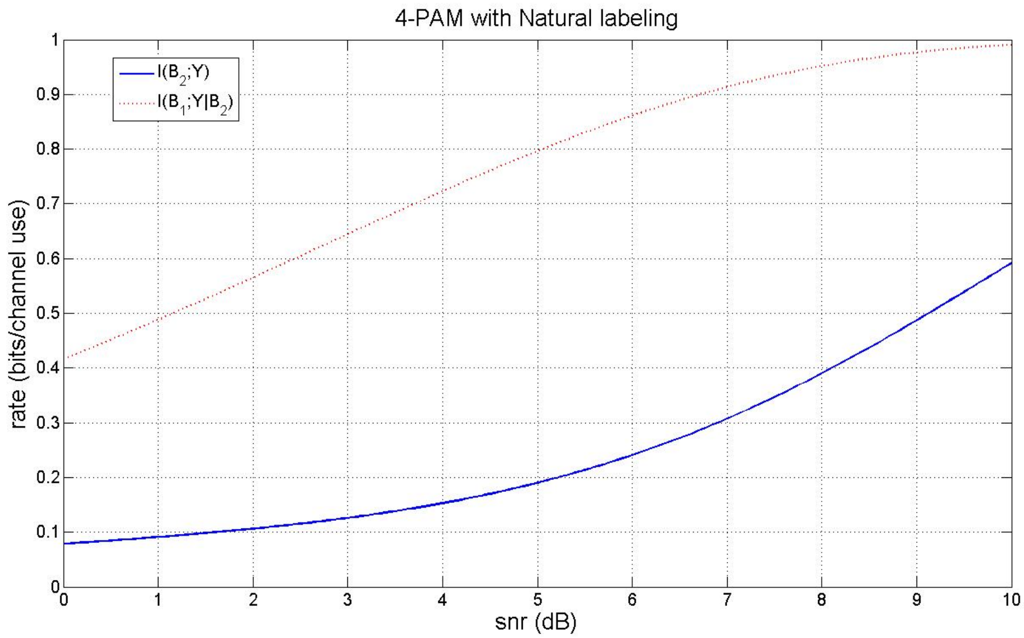

Figure 3.12: second order of decoding given in (3.168)

The significant difference in the rates achieved by the second bit (red graphs) in both cases can be easily understood form the constellation structure. $I(B_2; Y|B_1)$ and $I(B_1; Y|B_2)$ can be seen as a weighted sum of two mutual information whose values depend on the relative positioning of points in each subset determined by a specific bit. For example in the first case when $B_1 = 0$ the subset contains two points -3, -1 but in the second case when $B_2 = 0$ the subset contains -3 , +1 which are further apart compared to the fist case so

$$I(B_2; Y|B_1 = 0) < I(B1; Y|B_2 = 0). \tag{3.183}$$

The same argument can be applied to get

$$B_2; Y|B_1 = 1) < I(B1; Y|B_2 = 1) \tag{3.184}$$

Hence

$$I(B_2; Y|B_1) < I(B_1; Y|B_2). \tag{3.185}$$

Comparing the rate achieved by the first bit in each case reduces to comparing the two following terms:

$$\frac{1}{2}[h(Y|B_1 = 0) + h(Y|B_1 = 1)] = h(Y|B_1 = 0) \tag{3.186}$$

$$\frac{1}{2}[h(Y|B_2 = 0) + h(Y|B_2 = 1)] = h(Y|B_2 = 0) \tag{3.187}$$

In the second term, the two Gaussian components are twice spaced than the components in the first term, hence the second mixed-Gaussian term has more power. On the other hand, the first term is more Gaussian like distribution so the comparison needs to evaluate the exact values. But we already know from previous section that

$$h(Y|B_1 = 0) < h(Y|B_2 = 0) \tag{3.188}$$

therefore

$$I(Y; B_2) < I(Y; B1) \tag{3.189}$$

As you see from two above figures, each decoding order results in in totaly different R_1 and R_2 but the summation $R = R_1 + R_2$ is the same and does not depend on the order of decoding:

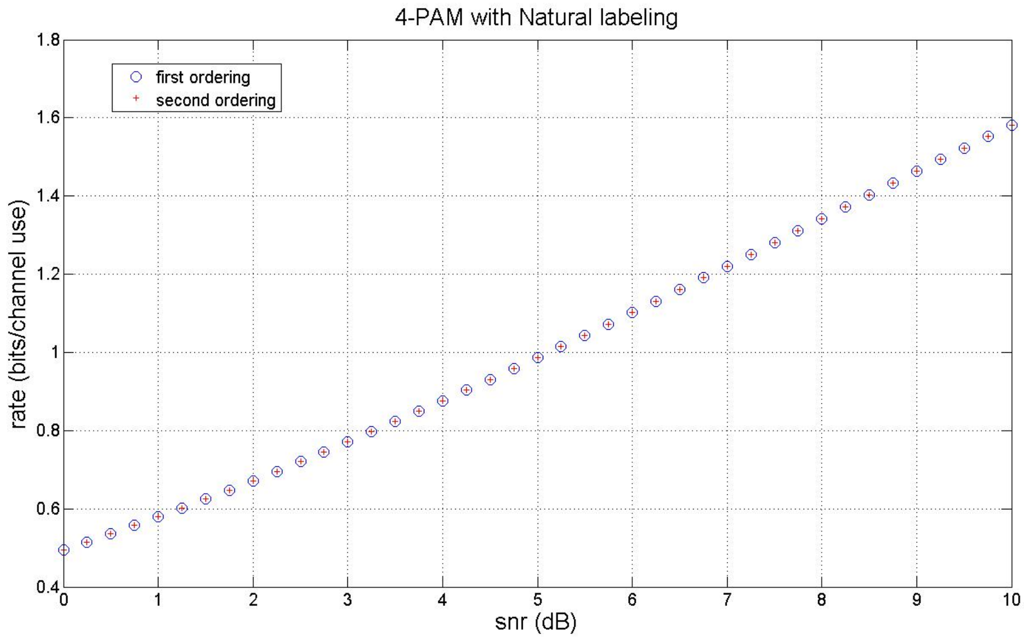

Figure 3.13: both order of decoding give the same rate

Chapter 4

Bit-Interleaved Coded Modulation

In Bit-Interleaved Coded Modulation (BICM) scheme, we introduce an interleaver between the binary encoder and the modulator. The interleaver tries to diminish the correlation between different bit positions. This approach makes the design of encoder and modulator independent from each other and we can design them separately to optimize the overall performance of the system. Each m-bit subsequent of the encoder output are mapped to a constellation point through a labelling function $\mu : \{0,1\}^m \rightarrow \mathcal{X}$. For this model, Caire et al. defined a BICM capacity [14] given by:

$$C_{\mathcal{X}}^{BICM} = \sum_{i=1}^{m} I(B_i; Y) \tag{4.1}$$

and in [15] using typical sequence,s the authors show that it is indeed achievable. In this case, in contrast to Coded Modulation scheme, the choice of labelling highly affects the performance of the system and the resulting rate.

Figure 4.1: BICM Transmitter

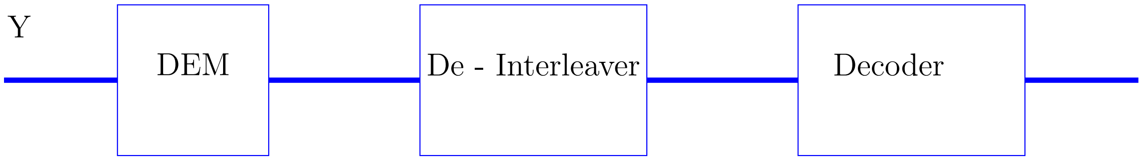

Figure 4.2: BICM Receiver

For a given constellation, we are trying to find the optimum binary labelling from a BICM capacity maximization point of view. In [14], the authors conjectured that Gray labelling maximizes BICM capacity but we show that it is not true for all range of snr. [8] compares Gray labelling with Set Partitioning labelling which is used in TCM (Trellis Coded Modulation) scheme for various PSK and QAM constellations.

In [16] optimal binary labelling, input distributions and input alphabet for BICM capacity are discussed at low snr regime. By some examples it shows that for different snr regions the BICM capacity is maximized with different labellings.

Calculating BICM capacity requires to compute m mutual information for each position. Because of the mixed-Gaussian noise term in each subchannel, we can not obtain a closed form expression for BICM capacity in general.

Hence, the problem here is to find parameter vectors $\mathbf{w} = (w_1, ..., w_m)$ to maximize $\sum_{i=1}^{m} I(B_i; Y)$ subject to underlying constraints of the problem. Here we assume average power constraint $E[X^2] \leq P$ which is equivalent to:

$$\sum_{i=1}^{m} w_i^2 \leq P \tag{4.2}$$

We formulate the following optimization problem:

$$\begin{aligned} \underset{\mathbf{w}}{\text{maximize}} \quad & \sum_{i=1}^{m} I(B_i; Y) \\ \text{subject to} \quad & \sum_{i=1}^{m} w_i^2 \leq P \\ & \mathbf{w} \succeq 0. \end{aligned} \tag{4.3}$$

This problem here is similar to (3.73) in nature with different objective function compared to (3.73).

There are different ways to solve this optimization problem in which the objective function doesn't admit a closed form expression. The objective function is not convex, hence, special care must be taken when trying to find global optimal using numerical methods.

4.1 Effect of Labelling

With the introduction of interleaver we can consider the BICM as m parallel binary input continues-output channel with rate $I(B_k; Y)$ associated to k th sub-channel. with the assumption of an interleaver with infinite depth, the k subchannels are considered to be independent [14] [25], or the correlation among sub-channels are assumed to be neglected. So the general rate achieved in this scheme is the summation of rates of sub-channels as given in (4.1). Using data processing inequality, [14] shows that BICM capacity for a given constellation is always less than or at most equal to Coded Modulation capacity of the constellation.

$$C_{\mathcal{X}}^{BICM} \leq C_{\mathcal{X}}^{CM} \tag{4.4}$$

I have calculated and plotted the BICM capacity for 4-PAM and 8-PAM constellations for a range of snr. For comparison purposes, I also plotted the CM capacity of both constellations on the same plot.

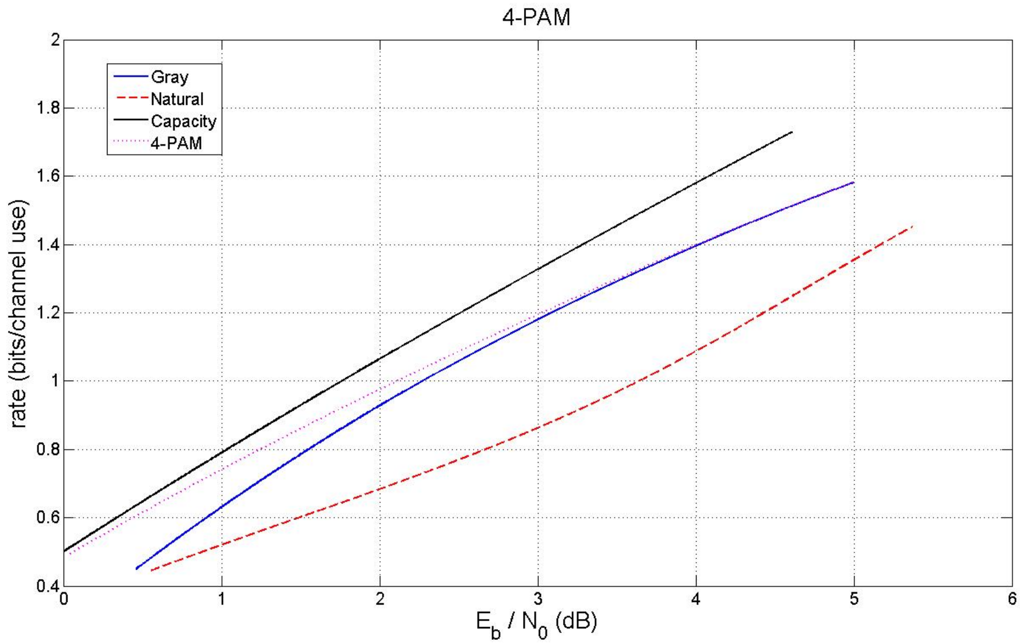

Figure 4.3: BICM capacity for two different labellings , 4-PAM

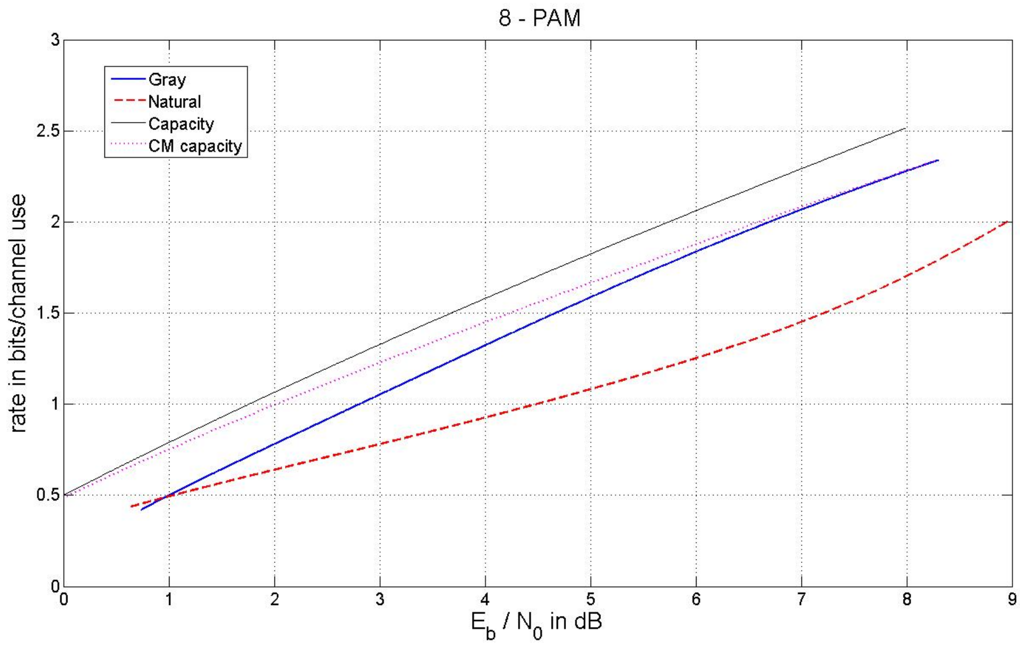

Figure 4.4: BICM capacity for two different labellings, 8-PAM(in Eb/N0)

As it is seen from both plots, the BICM capacity for both labellings are less than CM capacity. Gray labelling is optimal in high snr regime and also achieves CM capacity at high snr regime. Gray labelling has more than 1.5dB gain compared to Natural labelling for R = 1 bit/channel use.

There is a relatively huge gap between CM capacity and BICM capacity for Gary labelling at medium to low snr regime. This is where AWM comes to play its role. As it is seen from the following figure, it compensate most of the gap in this snr region.

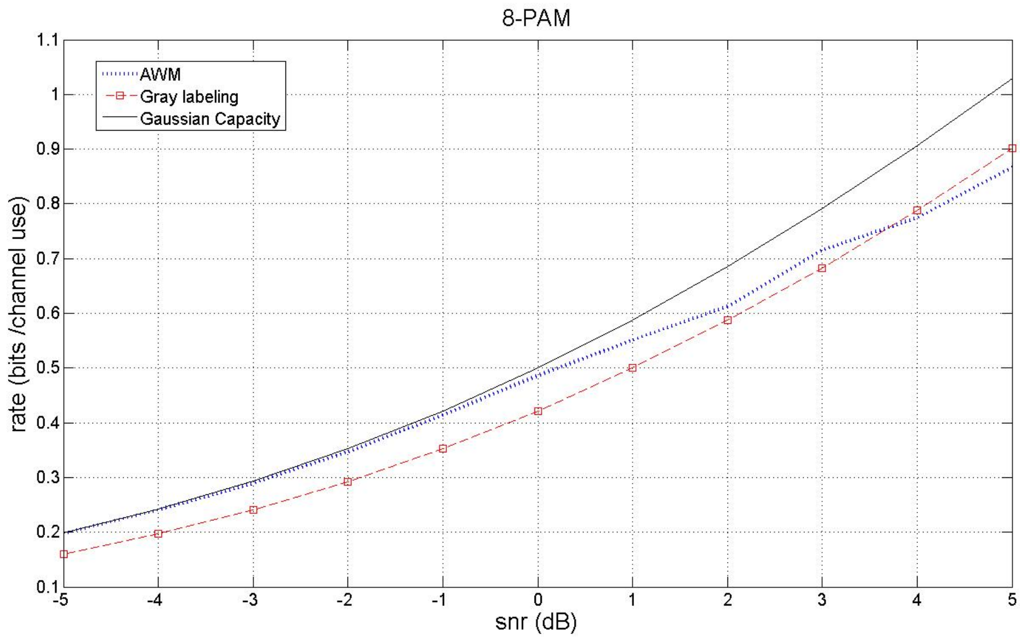

Figure 4.5: BICM capacity for two different labellings, 8-PAM

4.2 Asymptotic Behaviour of BICM Capacity

We can approximate any capacity function $C(snr)$ around snr $= 0$ using Taylor series expansion as

$$C(snr) = \alpha\ snr + \beta\ snr^2 + O(snr^3). \tag{4.5}$$

In the special case of our interest, the BICM capacity for a constellation is defined in [14] as

$$C^{BICM} = \sum_{i=1}^{m} I(Y; B_i). \tag{4.6}$$

We can express BICM capacity as CM capacity terms.

Lemma 3.

$$C^{BICM} = mI(Y;X) - \sum_{i=1}^{m} \sum_{b\ \in\{0,1\}} P_{B_i}(b)\ I(Y;X|B_i = b) \tag{4.7}$$

Proof: For each term in the summation

$$\begin{aligned} I(Y;B_i) &= h(Y) - h(Y|B_i) \\ &= h(Y) - h(Y|X) + h(Y|X) - h(Y|B_i) \\ &= I(Y;X) - I(Y;X|B_i) \end{aligned} \tag{4.8}$$

since $h(Y|X) = h(Y|X, B_i)$.

$$I(Y;X|B_i) = \sum_{b\ \in\{0,1\}} P_{B_i}(b)\ I(Y;X|B_i = b) \tag{4.9}$$

so

$$\begin{aligned} C^{BICM} &= \sum_{i=1}^{m} [\ I(Y;X) - \sum_{b\ \in\{0,1\}} P_{B_i}(b)\ I(Y;X|B_i = b)\ ] \\ &= mI(Y;X) - \sum_{i=1}^{m} \sum_{b\ \in\{0,1\}} P_{B_i}(b)\ I(Y;X|B_i = b) \end{aligned} \tag{4.10}$$

Each term in (4.7) is a Coded Modulation mutual information. Using Theeorem 7 in [**?**]

$$\alpha_{i,b}^{CM} = \log e\ [1 - \frac{(E[X|B_i = b])^2}{E[X^2|B_i = b]}] \tag{4.11}$$

So

$$\alpha^{BICM} = m\ \alpha^{CM} - \sum_{i=1}^{m} \sum_{b\in} P_{B_i}(b)\ \alpha_{i,b}^{CM} \tag{4.12}$$

calculating the second term of (4.12)

$$\begin{aligned} &\sum_{i=1}^{m} \sum_{b\ \in\{0,1\}} P_{B_i}(b) \log e\ (1 - \frac{(E[X|B_i = b])^2}{E[X^2|B_i = b]}) \\ &= \log e\ [m - \sum_{i=1}^{m} \sum_{b\ \in\{0,1\}} P_{B_i}(b)\ \frac{(E[X|B_i = b])^2}{E[X^2|B_i = b]} \end{aligned} \tag{4.13}$$

The first term of (4.12) is

$$m\ \log e\ (1 - \frac{(E[X])^2}{E[X^2]}) = m\ loge - m \log e\ \frac{(E[X])^2}{E[X^2]}. \tag{4.14}$$

Hence the first coefficient of the Taylor expansion for BICM scheme is

$$\alpha^{BICM} = \log e\ \Big[\sum_{i=1}^{m} \sum_{b\ \in\{0,1\}} P_{B_i}(b)\ \frac{(E[X|B_i = b])^2}{E[X^2|B_i = b]} \quad - \quad m\ \frac{(E[X])^2}{E[X^2]}\Big]. \tag{4.15}$$

In the case

$$E[X] = 0 \tag{4.16}$$

$$P_{B_i}(0) = P_{B_i}(1) = \frac{1}{2} \tag{4.17}$$

We can simplify (4.15) to get

$$\alpha^{BICM} = \frac{\log e}{2} \sum_{i=1}^{m} \sum_{b\ \in\{0,1\}} \frac{(E[X|B_i = b])^2}{E[X^2|B_i = b]}. \tag{4.18}$$

In particular for our model with its underlying labelling

$$\begin{aligned} E[X|B_1 = 0] &= -w_1 \\ E[X|B_1 = 1] &= w_1 \\ E[X|B_2 = 0] &= -w_2 \\ E[X|B_2 = 1] &= w_2 \end{aligned} \tag{4.19}$$

$$E[X^2|B_i = b] = w_1^2 + w_2^2 \tag{4.20}$$

plugging in (4.18) yields

$$\alpha^{BICM} = \frac{\log e}{2(w_1^2 + w_2^2)}\ 2(w_1^2 + w_2^2) = \log e. \tag{4.21}$$

If we use the constellation points produced by the model but with Gray labelling

$$
\begin{aligned}
E[X|B_1 = 0] &= -w_1 \\
E[X|B_1 = 1] &= w_1 \\
E[X|B_2 = 0] &= 0 \\
E[X|B_2 = 1] &= 0
\end{aligned} \tag{4.22}
$$

$$E[X^2|B_1 = 0] = w_1^2 + w_2^2 \tag{4.23}$$

$$E[X^2|B_1 = 1] = w_!^2 + w_2^2 \tag{4.24}$$

$$E[X^2|B_2 = 0] = (w_1 + w_2)^2 \tag{4.25}$$

$$E[X^2|B_2 = 1] = (w_1 - w_2)^2 \tag{4.26}$$

$$
\begin{aligned}
\alpha^{BICM} &= \frac{\log e}{2}\,[\frac{w_1^2}{w_1^2 + w_2^2} + \frac{w_1^2}{w_1^2 + w_2^2}] \\
&= \log e\,\frac{w_1^2}{w_1^2 + w_2^2} < \log e
\end{aligned} \tag{4.27}
$$

4.3 Higher Order Optimality

In [26], the coefficients of Taylor expansion for a coded modulation capacity of a general constellation is given. We expand capacity (in nats per channel use) in the form of

$$C = \alpha snr + \beta snr^2 + O(snr^2). \tag{4.28}$$

Using Theorem 5 in [26] we will obtain simple expressions for α and β in Coded Modulation scheme and then we will extend the results for BICM scheme. In the following, we restate Theorem 5 in [26]

Theorem 8. *let $x = (x_1, \cdots, x_k)$ and $z = (z_1, \cdots, z_k)$ be two independent k-dimensional real-valued random vector where the components of z are are iid Gaussian random variables with parameters $(0, \sigma^2)$. If there exists constant δ and μ such that*

$$E[|x|^{4+\delta}] \leq (\log \frac{1}{\sqrt{snr}})^{\mu} \tag{4.29}$$

then

$$\begin{aligned} I(x\ ;\ \sqrt{snr}x + z) &= \frac{\log e}{2\sigma^2}\ Tr\{cov(x)\} \\ &- \frac{\log e}{4\sigma^2}\ Tr\{cov^2(x)\} + O(snr^2) \end{aligned} \tag{4.30}$$

Proof: see [26]

In one-dimensional signalling, we have

$$Y = \sqrt{snr}X + N \tag{4.31}$$

Where X and N are independent and N is a Gaussian random variable with zero mean and $\sigma^2 = 1$. Simplifying the above theorem for this one-dimensional constellation case, we get

$$\begin{aligned} I(X\ ;\ \sqrt{snr}X + N) &= \frac{\log e}{2 \times 1}\ var(X)\ snr \\ &- \frac{\log e}{4 \times 1}\ var^2(X)\ snr^2 + O(snr^2). \end{aligned} \tag{4.32}$$

so

$$\alpha = \frac{\log e}{2}\ var(X) \tag{4.33}$$

$$\beta = -\frac{\log e}{4}\ var^2(X) \tag{4.34}$$

Using (4.7) we can obtain expressions for the first two coefficients of Taylor expansion in BICM scheme.

$$C^{BICM} = mI(Y;X) - \sum_{i=1}^{m} \sum_{b\ \in\{0,1\}} P_{B_i}(b)\ I(Y;X|B_i = b) \tag{4.35}$$

Following common notation in BICM literature, we denote the subset of constellation points which have bit b in position i by $\mathcal{X}_b^i$.

According to (4.7) the corresponding coefficients for BICM capacity are

$$\alpha^{BICM} = m\ \alpha^{CM} - \sum_{i=1}^{m} \sum_{b\in\{0,1\}} P_{B_i}(b)\ \alpha^{CM}(\mathcal{X}_b^i) \tag{4.36}$$

$$\beta^{BICM} = m\ \beta^{CM} - \sum_{i=1}^{m} \sum_{b\in\{0,1\}} P_{B_i}(b)\ \beta^{CM}(\mathcal{X}_b^i) \tag{4.37}$$

We bring all the pieces together in the following theorem.

Theorem 9. *For BICM with equi-probable binary encoding the coefficients for the first two terms in Taylor expansion around snr = 0 are given as*

$$\alpha = \frac{\log e}{2}\ \Big[m\ var(X)\ -\ \frac{1}{2} \sum_{i=1}^{m} \sum_{b\in\{0,1\}} var(\mathcal{X}_b^i)\ \Big] \tag{4.38}$$

$$\beta = \frac{\log e}{4}\ \Big[\frac{1}{2} \sum_{i=1}^{m} \sum_{b\in\{0,1\}} var^2(\mathcal{X}_b^i)\ -\ m\ var^2(X)\ \Big] \tag{4.39}$$

Now, we apply (4.39) to our model. For m = 2 we have

$$\mathcal{X}_0^1 = \{\ -w_1 - w_2,\quad -w_1 + w_2\ \} \tag{4.40}$$

$$\mathcal{X}_1^1 = \{\ w_1 - w_2,\quad w_1 + w_2\ \} \tag{4.41}$$

$$\mathcal{X}_0^2 = \{\ -w_1 - w_2,\quad w_1 - w_2\ \} \tag{4.42}$$

$$\mathcal{X}_1^2 = \{\ -w_1 + w_2,\quad -w_1 + w_2\ \} \tag{4.43}$$

Therefore

$$var(\mathcal{X}_0^1) = w_2^2 \tag{4.44}$$

$$var(\mathcal{X}_1^1) = w_2^2 \tag{4.45}$$

$$var(\mathcal{X}_0^2) = w_1^2 \tag{4.46}$$

$$var(\mathcal{X}_1^2) = w_1^2 \tag{4.47}$$

So

$$\begin{aligned}\alpha &= \frac{\log e}{2}\,\left[2(w_1^2+w_2^2)-\frac{1}{2}(2w_1^2+2w_2^2)\right] \\ &= \frac{\log e}{2}(w_1^2+w_2^2) = \frac{\log e}{2}\,var(X) = \alpha^{CM} \end{aligned} \tag{4.48}$$

$$\begin{aligned}\beta &= \frac{\log e}{4}\,\left[\frac{1}{2}(2w_1^4+2w_2^4)-2(w_1^2+w_2^2)^2\right] \\ &= -\frac{\log e}{4}\,\left[2(w_1^4+w_2^4+2w_1^2w_2^2)-w_1^4-w_2^4\right] = \\ &= -\frac{\log e}{4}\,[(w_1^2+w_2^2)^2+2w_1^2w_2^2] \\ &= -\frac{\log e}{4}\,[var^2(X)+2w_1^2w_2^2] \\ &< -\frac{\log e}{4}\,var^2(X) = \beta^{CM} \end{aligned} \tag{4.49}$$

So

$$\alpha^{AWM} = \alpha^{CM} \tag{4.50}$$

$$\beta^{AWM} < \beta^{CM} \tag{4.51}$$

but if we use the constellation generated according to (2.5) with Gray labelling we will obtain:

$$\mathcal{X}_0^1 = \{\ -w_1-w_2, \quad -w_1+w_2\ \} \tag{4.52}$$

$$\mathcal{X}_1^1 = \{\ w_1-w_2, \quad w_1+w_2\ \} \tag{4.53}$$

$$\mathcal{X}_0^2 = \{\ -w_1-w_2, \quad w_1+w_2\ \} \tag{4.54}$$

$$\mathcal{X}_1^2 = \{\ -w_1+w_2, \quad w_1-w_2\ \} \tag{4.55}$$

Therefore

$$var(\mathcal{X}_0^1) = w_2^2 \tag{4.56}$$

$$var(\mathcal{X}_1^1) = w_2^2 \tag{4.57}$$

$$var(\mathcal{X}_0^2) = (w_1+w_2)^2 \tag{4.58}$$

$$var(\mathcal{X}_1^2) = (w_1-w_2)^2 \tag{4.59}$$

$$
\begin{aligned}
\alpha &= \frac{\log e}{2}[2(w_1^2 + w_2^2) - \frac{1}{2}(2w_2^2 + (w_1 - w_2)^2 + (w_1 + w_2)^2] \\
&= \frac{\log e}{2}\ [w_1^2 + w_2^2 - w_2^2] = \frac{\log e}{2}\ w_1^2 < \frac{\log e}{2}\ (w_1^2 + w_2^2) = \alpha^{CM}
\end{aligned}
\tag{4.60}
$$

$$
\begin{aligned}
\beta &= -\frac{\log e}{4}\ [\frac{1}{2}(2w_2^4 + (w_1 - w_2)^4 + (w_1 + w_2)^4) - 2(w_1^2 + w_2^2)^2] \\
&= -\frac{\log e}{4}\ [2w_1^2 w_2^2 - w_1^4] > -\frac{\log e}{4}\ [2w_1^2 w_2^2 + w_1^4 + w_2^4] \\
&= -\frac{\log e}{4}\ var^2(X) = \beta^{CM}
\end{aligned}
\tag{4.61}
$$

So

$$\alpha^{Gray} < \alpha^{CM} \tag{4.62}$$

$$\beta^{Gray} > \beta^{CM} \tag{4.63}$$

Which shows sub-optimality of Gray labelling in low snr. For 4-PAM with Natural labelling:

$$\mathcal{X}_0^1 = \{-3\mu\ , \quad -\mu\ \} \tag{4.64}$$

$$\mathcal{X}_1^1 = \{+\mu\ , \quad +3\mu\} \tag{4.65}$$

$$\mathcal{X}_0^2 = \{-3\mu, \quad +3\mu\ \} \tag{4.66}$$

$$\mathcal{X}_1^2 = \{-\mu, \quad +\mu\ \} \tag{4.67}$$

$$var(X) = 5\mu^2 \tag{4.68}$$

$$var(\mathcal{X}_0^1) = \mu^2 \tag{4.69}$$

$$var(\mathcal{X}_1^1) = \mu^2 \tag{4.70}$$

$$var(\mathcal{X}_0^2) = 9\mu^2 \tag{4.71}$$

$$var(\mathcal{X}_1^2) = \mu^2 \tag{4.72}$$

$$\begin{aligned}
\alpha &= \frac{\log e}{2}[10\mu^2 - \frac{1}{2}(12\mu^2)] \\
&= \frac{\log e}{2}\ 4\mu^2 \\
&= \frac{4}{5}(\frac{\log e}{2}\ var(X)) < \alpha^{CM} \qquad (4.73)
\end{aligned}$$

$$\begin{aligned}
\beta &= \frac{\log e}{4}\ [\frac{1}{2}(3\mu^4 + 81\mu^4) - 2(25\mu^4)] \qquad (4.74) \\
&= -\frac{\log e}{4}[12\mu^4] = \frac{12}{25}\beta^{CM} > \beta^{CM} \qquad (4.75)
\end{aligned}$$

4-PAM with Natural labelling is not first-order optimal.

4.3.1 Optimal Labelling Criteria

we are trying to find labelling for a given constellation to make BICM capacity as close as possible to CM capacity. for small snr values, this is equivalent to make the first few terms in Taylor expansion similar to corresponding terms in CM expansion.

$$\alpha^{BICM} = \alpha^{CM} \qquad (4.76)$$

$$\frac{\log e}{2}\ var(X) = \frac{\log e}{2}\ \Big[m\ var(X)\ -\ \frac{1}{2}\sum_{i=1}^{m}\sum_{b\in\{0,1\}} var(\mathcal{X}_b^i)\ \Big] \qquad (4.77)$$

or

$$(m-1)\ var(X) = \frac{1}{2}\sum_{i=1}^{m}\sum_{b\in\{0,1\}} var(\mathcal{X}_b^i) \qquad (4.78)$$

and trying to design labelling in such a way to minimize the difference in the second term.

$$\begin{aligned}
\beta^{CM} - \beta^{BICM} &= -\frac{\log e}{4}\ var^2(X) \\
&\quad - \frac{\log e}{4}\ \Big[\frac{1}{2}\sum_{i=1}^{m}\sum_{b\in\{0,1\}} var^2(\mathcal{X}_b^i)\ -\ m\ var^2(X)\ \Big] \\
&= \frac{\log e}{4}\ \Big[(m-1)\ var^2(X) - \frac{1}{2}\sum_{i=1}^{m}\sum_{b\in\{0,1\}} var^2(\mathcal{X}_b^i)\Big] \qquad (4.79)
\end{aligned}$$

References

[1] Feng-Wen Sun and H.C.A. van Tilborg. Approaching capacity by equiprobable signaling on the gaussian channel. *Information Theory, IEEE Transactions on*, 39(5):1714 –1716, sep 1993.

[2] K. Ruotsalainen. On the construction of the higher dimensional constellations. In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, page 490, 2002.

[3] Jr. Forney, G., R. Gallager, G. Lang, F. Longstaff, and S. Qureshi. Efficient modulation for band-limited channels. *Selected Areas in Communications, IEEE Journal on*, 2(5):632 – 647, sep 1984.

[4] Jr. Forney, G.D. Multidimensional constellations. ii. voronoi constellations. *Selected Areas in Communications, IEEE Journal on*, 7(6):941 –958, aug 1989.

[5] Jr. Forney, G.D. and L.-F. Wei. Multidimensional constellations. i. introduction, figures of merit, and generalized cross constellations. *Selected Areas in Communications, IEEE Journal on*, 7(6):877 –892, aug 1989.

[6] G. Foschini, R. Gitlin, and S. Weinstein. Optimization of two-dimensional signal constellations in the presence of gaussian noise. *Communications, IEEE Transactions on*, 22(1):28 – 38, jan 1974.

[7] Maged F. Barsoum, Christopher Jones, and Michael Fitz. Constellation design via capacity maximization. In *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, pages 1821 –1825, june 2007.

[8] G. Ungerboeck. Channel coding with multilevel/phase signals. *Information Theory, IEEE Transactions on*, 28(1):55 – 67, jan 1982.

[9] L.H. Ozarow and A.D. Wyner. On the capacity of the gaussian channel with a finite number of input levels. *Information Theory, IEEE Transactions on*, 36(6):1426 –1428, nov 1990.

[10] H. Imai and S. Hirakawa. A new multilevel coding method using error-correcting codes. *Information Theory, IEEE Transactions on*, 23(3):371 – 377, may 1977.

[11] J. Huber and U. Wachsmann. Capacities of equivalent channels in multilevel coding schemes. *Electronics Letters*, 30(7):557 –558, mar 1994.

[12] Udo Wachsmann, Lehrstuhl Fur Nachrichtentechnik, and Johannes Huber. Power and bandwidth efficient digital communication using turbo codes in multilevel codes. *Europ. Trans. Telecommun. (ETT*, 6:557–567, 1995.

[13] Y. Kofman, E. Zehavi, and S. Shamai. Performance analysis of a multilevel coded modulation system. *Communications, IEEE Transactions on*, 42(234):299 –312, feb/mar/apr 1994.

[14] G.Taricco G.Caire and E.Biglieri. Bit-interleaved coded modulation. *IEEE Transactions on Information Theory*, 44:927–946, 1998.

[15] A. Martinez, A. Guillen i Fabregas, G. Caire, and F. Willems. Bit-interleaved coded modulation revisited: A mismatched decoding perspective. *Information Theory, IEEE Transactions on*, 55(6):2756 –2765, june 2009.

[16] E. Agrell and A. Alvarado. Optimal alphabets and binary labelings for bicm at low snr. *Information Theory, IEEE Transactions on*, 57(10):6650 –6672, oct. 2011.

[17] C. E. SHANNON. A mathematical theory of communication. *IEEE Transactions on Information Theory*, 45:1361–1391, 1999.

[18] T. Cover and J.Thomas. *Elements of Information Theory*. Wiley, Hoboken, NJ, 2006.

[19] J. Huang and S.P. Meyn. Characterization and computation of optimal distributions for channel coding. *Information Theory, IEEE Transactions on*, 51(7):2336 –2351, july 2005.

[20] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Hoboken, NJ, 2004.

[21] D. Guo, S. Shamai, and S. Verdu. Mutual information and minimum mean-square error in gaussian channels. *Information Theory, IEEE Transactions on*, 51(4):1261 –1282, april 2005.

[22] Yihong Wu and S. Verdu and. The impact of constellation cardinality on gaussian channel capacity. In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pages 620 –628, 29 2010-oct. 1 2010.

[23] N. Varnica, Xiao Ma, and A. Kavcic. Capacity of power constrained memoryless awgn channels with fixed input constellations. In *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, volume 2, pages 1339 – 1343 vol.2, nov. 2002.

[24] A. Abedi, M.E. Thompson, and A.K. Khandani. Application of cumulant method in performance evaluation of turbo-like codes. In *Communications, 2007. ICC '07. IEEE International Conference on*, pages 986 –989, june 2007.

[25] U. Wachsmann, R.F.H. Fischer, and J.B. Huber. Multilevel codes: theoretical concepts and practical design rules. *Information Theory, IEEE Transactions on*, 45(5):1361 – 1391, jul 1999.

[26] V.V. Prelov and S. Verdu. Second-order asymptotics of mutual information. *Information Theory, IEEE Transactions on*, 50(8):1567 – 1580, aug. 2004.

[27] Yihong Wu and S. Verdu. Functional properties of minimum mean-square error and mutual information. *Information Theory, IEEE Transactions on*, 58(3):1289 –1301, march 2012.

[28] I.Fabregas A. Martinez, A.Guillen and G.Caire. Bit-interleaved coded modulation. *Foundations and Trends in Commununication and Information Theory*, 5:1–153, 2008.

[29] A. Martinez, A. Guillen i Fabregas, G. Caire, and F. Willems. Bit-interleaved coded modulation in the wideband regime. *Information Theory, IEEE Transactions on*, 54(12):5447 –5455, dec. 2008.

[30] A. Guille andn i Fabregas and A. Martinez. Bit-interleaved coded modulation with shaping. In *Information Theory Workshop (ITW), 2010 IEEE*, pages 1 –5, 30 2010-sept. 3 2010.

[31] M.F. Huber, T. Bailey, H. Durrant-Whyte, and U.D. Hanebeck. On entropy approximation for gaussian mixture random vectors. In *Multisensor Fusion and Integration for Intelligent Systems, 2008. MFI 2008. IEEE International Conference on*, pages 181 –188, aug. 2008.

[32] M.F. Huber and U.D. Hanebeck. Progressive gaussian mixture reduction. In *Information Fusion, 2008 11th International Conference on*, pages 1 –8, 30 2008-july 3 2008.

[33] D. Schieferdecker and M.F. Huber. Gaussian mixture reduction via clustering. In *Information Fusion, 2009. FUSION '09. 12th International Conference on*, pages 1536 –1543, july 2009.

[34] A.R. Runnalls. Kullback-leibler approach to gaussian mixture reduction. *Aerospace and Electronic Systems, IEEE Transactions on*, 43(3):989 –999, july 2007.